universität wien

Diplomarbeit

Titel

Validation of the ‘Viennese Emotional Test Battery’ -

Validating the items of the Viennese Emotional Go/no-go Task, Viennese Emotional Stroop Task and the Viennese Affective Flexibility Task

verfasst von

Matthias Deckert

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Wien, 2014

Studienkennzahl:	A 298
Studienrichtung:	Diplomstudium Psychologie
Betreuerin:	Univ. Prof. Mag. Dr. Ulrike Willinger

Acknowledgment

First of all I would like to thank Univ.-Prof.-Mag. Dr. Ulrike Willinger for inspiring me to wake my enthusiasm for emotion research and for giving me the chance to write this diploma thesis. Furthermore I would like to thank her for the huge amount of support she provided me as well as for every single idea for improvement.

I would also like to thank Dr. Michaela Schmöger for all the great support she gave me and for always helping me out with the programs E-Prime 2.0 and SPSS.

I would like to thank my parents for giving me the chance to study psychology and for their great emotional and financial support throughout all these years.

To a very special person I simply want to say: Thank you for everything.

Last but not least I thank every single person who participated in this diploma thesis, without you I wouldn’t sit here being busy writing an acknowledgment.

Table of Contents

Theoretical background

1. The Viennese Emotional Test Battery – Description of the Viennese Emotional Stroop Task, Viennese Affective Flexibility Task and Viennese Emotional Go/no-go task

In the years 2012 and 2013 Willinger, Schmoeger, Deckert and Auff designed the 'Viennese Emotional Test Battery' (VETeBa), a test battery for measuring different aspects of affective information processing. VETeBa consists of three tasks each one designed according to different information processing paradigms and presenting different sorts of emotional stimuli in order to assess affective information processing from different point of views. In the following sections each of the three tasks of the VETeBa namely the 'Viennese Emotional Stroop Task', the 'Viennese Affective Flexibility Task' and the 'Viennese Emotional Go/no-go Task' will be described.

1.1 The Viennese Emotional Stroop Task

In 2012 Willinger, Schmoeger, Deckert and Auff designed the ‚Viennese Emotional Stroop Task' in the style of Etkin and colleagues (Etkin, Egner, Peraza, Kandel, & Hirsch, 2006), presenting fearful as well as happy faces with the emotion words 'fear' or 'happiness' written across the face. The task was designed comprising of two different conditions namely the 'face-condition' and the 'word-condition' which assess different requirements regarding affective information processing. In the 'face-condition' participants have to react to the facially expressed emotions 'happiness' or 'fear' via button press whilst ignoring the emotion words ('happiness' or 'fear') written across the face with red letters. In the 'word-condition' participants have to react to the emotion words 'happiness' or 'fear' via button press whilst ignoring the underlying facial expressions (happy or fearful faces). In order to fulfill the requirement of the Stroop paradigm the presented stimuli could be congruent in nature (happy faces with word 'happiness' as well as fearful faces with word 'fear') or incongruent (happy faces with word 'fear' as well as fearful faces with word 'happiness'). For both conditions the

same stimuli are used albeit different functions of processing are required. Depending on the requirements of the current condition participants have to ignore the emotional properties of either faces or words in order to react adequately.

For the Viennese Emotional Stroop Task human faces displaying fearful as well as happy expressions were randomly drawn from the Karolinska Directed Emotional Faces Database (KDEF - Lundqvist et al. 1998) each face showing a straight gaze. In the style of Etkin and colleagues (2006) faces were cropped and presented in different shades of grey in order to attenuate the confounding factor of facial complexion. Furthermore this was done so as to contrast the emotion word (written in red letters across the face) from the face in a more elaborate way.

1.2 **The Viennese Affective Flexibility Task**

The Viennese Affective Flexibility Task was designed in 2013 by Willinger, Schmoeger, Deckert and Auff in the style of Barcelo (2003). In his computerized 'Madrid Card Sorting Test' participants are presented four basic cards differing in different dimensions (color, form, and number of presented objects) whilst subsequently they were presented single cards below which they had to sort to one of the basic cards according to one of the dimensions. Regularly albeit after different numbers of trials the predominant sorting principle changed and participants had to find out the new sorting principle. In the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013) similarly four basic cards are presented in a computerized task albeit presenting different stimuli properties and thus different sorting categories. The four basic cards present faces in different numbers (one - four faces) with different emotions (fear, sadness, anger and happiness) surrounded by different frame colors (blue, red, yellow, and green). Participants are subsequently presented single cards below which they have to sort to one of these basic cards according to one of the sorting dimensions color (frame), number or displayed emotion. From time to time the predominant sorting principle changes abruptly and participants have to find out which sorting principle is to use then. Likewise to the 'Madrid Card Sorting Test' (Barcelo, 2003) also the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013) requires successful execution of shifting, updating and rehearsal processes with the difference that also emotional cues have to be heeded.

Likewise to the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012) faces presenting angry, fearful, sad and happy straight gaze expressions were randomly drawn from the Karolinska Directed Emotional Faces Database (KDEF - Lundqvist et al., 1998) in order to be used for the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013).

1.3 **The Viennese Emotional Go/no-go Task**

In 2012 Willinger, Schmoeger, Deckert and Auff designed the Viennese Emotional Go/no-go task in the style of Gole and colleagues (Gole, Koechel, Schaefer, & Schienle, 2012). In this emotional go/no-go paradigm task participants are presented negative as well as neutral words to which they have to react to (go words) or to inhibit a reaction (no-go words) depending on the current condition. Comprising of eight conditions in sum each condition requires different reactions according to the valence of the words (semantic conditions - negative or neutral words), font (syntactic conditions - only first letter being capital or whole word being written in block capital) or combinations of both (mixed semantic and syntactic conditions).

The negative and neutral words used in the Viennese Emotional Go/no-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) were drawn from the Berlin Affective Word List Reloaded (BAWL-R) (Voe, et al., 2009).

1.4 **Implications for the validation of the Viennese Emotional Test Battery Items**

In the previous sections short descriptions of the VETeBa tasks and its items were presented showing that these tasks rely on stimuli consisting of either pictures of human faces displaying different emotions as well as on verbal material with different emotional connotations. These pictures as well as the words were drawn from standardized face- and word databases evaluated in the course of validation studies including samples from different countries (partly consisting of unknown nationalities and not offering information about native language – as will be described in the following chapters). Due to the unclear body of acquired knowledge concerning universalities between cultures regarding the perception and processing of emotional cues the clear assumption that the stimuli chosen for the VETeBa are equally perceived and processed by Austrians or German speaking people living in Austria

can’t be drawn. In order to be able to use the VETeBa tasks on samples consisting of Austrians as well as German speaking people living in Austria and to draw meaningful conclusions from the results obtained by the VETeBa tasks the necessity to validate the chosen stimuli within representative samples is implied. Due to the fact that the VETeBa tasks will be used on people of different gender and age groups this diploma thesis will additionally focus how age and gender influence the perception and processing of the VETeBa task items.

In the following chapters the necessity of validating the pictorial and verbal material chosen for the Viennese Emotional Test Battery tasks according to cultural, age- and gender aspects will be discussed extensively.

2. The perception and processing of human faces and human facially expressed emotions

2.1 Universality and specificity in facial emotion expression and recognition

In emotion research human facially expressed emotions are considered to be a core element in the experience and perception of emotion (Ekman, 1993; Izard, 1990). In comparison to other objects human faces seem to represent an own class in visual perception which seem to be recognized and processed in distinct ways and cortical areas (Wallis, 2013). In the course of the technical progress concerning research methods it has more and more become possible to analyze for similarities and differences in facial perception and processing. In the course of this progress it could be shown that there are cultural differences in the perception and processing of human faces (Busche, 2014; Cao, Wang, Rao, & Fu, 2013; Hills & Lewis, 2011; Ito, Masuda, & Man Wai Li, 2013; Ko, Lee, Yoon, Kwon, & Mather, 2011; Macchi Cassia, Luo, Pisacane, Li, & Lee, 2014; Miellet, Vizioli, He, Zhou, & Caldara, 2013; Senholzi & Ito, 2013; Tortosa, Lupianez, & Ruz, 2013) as well as cultural differences in the activation of specific brain regions underlying such processes (Derntl et al., 2012; Ding, Fu, & Lee, 2014).

In his publications Paul Ekman indicates the existence of a set of facial „basic emotions“, consisting of anger, sadness, fear, disgust, surprise, happiness and contempt, which seem to be

most consistently cross-culturally recognized (1972, 1992). Due to issues of choice and composition of samples, ecological validity of stimuli, goal of the study (looking for similarities instead of differences) and other methodological reasons, this universality has been questioned (for an overview see Elfenbein & Ambady, 2002; Scherer, Clark-Polner, & Mortillaro, 2012). Russel (1994) indicates that after controlling for such methodological aspects there should be a culture-specific perception of these emotions and that only a classification by the dimensions pleasure (valence) and (inner) arousal could be universal. Also Lang, Greenwald, Bradley and Hamm (1993) indicate valence and arousal to be fundamental dimensions of emotion. Substantiating these doubts on universality of emotion expression and emotion processing a number of recent studies show cultural differences in processing of facially expressed emotions (Engelmann & Pogosyan, 2013; Gendron, Roberson, van der Vyver, & Barrett, 2014; Gul & Humphreys, 2014; Jack, Garrod, Yu, Caldara, & Schyns, 2012; Stanley, Zhang, Fung, & Isaacowitz, 2013; Thibault, Levesque, Gosselin, & Hess, 2012; Vogel, Monesson, & Scott, 2012; Zhu, Ho, & Bonanno, 2013).

Ekman (1999) as well as Ekman and Matsumoto (2011) note concerning this debate about universality that there is strong evidence for these emotions having universal facial expressions in literate and preliterate cultures. Elfenbein and Ambady (2002) show in a meta-analysis that these previously described facial emotion expressions are universally recognized at better-than-chance levels, also Scherer and colleagues (2011) come to this conclusion. Similar to the findings of cultural differences in facial emotion perception also recent evidence sustaining the universality theory by Ekman (1972, 1992, 1999) could be found showing no cultural differences in facial emotion recognition (Andrade, Abreu, Duran, Veloso, & Moreira, 2013; Anguas-Wong & Matsumoto, 2007; Kayyal & Russell, 2013a; Ma, Li, Niu, Yu, & Yang, 2013; Matsumoto, Olide, & Willingham, 2009).

In search for possible reasons for cross-cultural differences Elfenbein and Ambady (2002) name the term „in-group advantage“. They describe it in ways that emotion perception is biased when the expressing person and the perceiving person are of the same cultural group, thus that the recognition accuracy is higher. Focusing on emotion perception in general Elfenbein and Ambady (2002) sum that concerning in-group advantages a number of incongruous results can be found. For facially expressed emotions Ekman (1972) for example couldn’t find an in-group advantage benefitting American or Japanese participants. Also Beaupré and Hess (2005) found no support for an in-group advantage whereas Biehl et al.

(1997) found evidence for a partial in-group advantage within high levels of agreement in general. They show that for emotions happy and surprise no cultural differences can be found, whereas for the exact level of agreement for sadness, disgust, contempt, anger and fear culture-specific aspects can be found.

Elfenbein and Ambady (2002) indicate in their meta-analysis that there is an in-group advantage between cultures and subcultures that accounts for the perception of facial expressed emotions but differs in magnitude (with contempt and happy having the smallest and disgust and fear having the greatest advantage). Scherer and colleagues (2011) show that there are cultural differences in encoding (expression) - decoding (recognition) patterns between Western and Non-Western cultures benefitting Western participants viewing Western photo models but maybe more interestingly also Non-Western participants viewing Western photo models. They show that for Non-Western participants the decoding of fear, sadness, surprise and disgust is more accurate when displayed by Western photo models rather than by Non-Western ones. They think that these results could (only) partly be explained by the fact that the Western stimulus photos are of superior quality because of the long Western research tradition. They show that overall no major differences in recognizing specific emotions can be found due to the origin of expressing and recognizing participants (Scherer et al., 2011). Elfenbein, Beaupre, Levesque and Hess (2007) on the other hand constitute within the dialectic theory framework that facially expressed emotions show cultural variations that can be seen as being similar to linguistic dialects. They show that there are subtle differences in activation of facial muscles, operationalized through the Facial Action Coding System (Ekman & Friesen 1978; Ekman, Friesen, & Hager, 2002), for several emotions but not for fear or disgust. A recent review (Elfenbein, 2013) summarizes empirical support that sustains the dialectic theory as well as the existence of in-group advantages concerning facial emotion recognition whilst other studies specify such in-group advantages as being some sort of own-race bias (Macchi Cassia et al., 2014) or being caused rather by culture than by race (Wickline, Bailey, & Nowicki, 2009).

Also recent studies concerning the recognition of facially expressed emotions show inconsistent results concerning this universality-specificity debate. Ma and colleagues (2013) show that Chinese show similar results to Western people when rating Western photo models. Likewise Andrade and colleagues (2013) show that ratings of North American photo models by Brazilians and North Americans show high agreement levels as well as Anguas-Wong and

Matsumoto (2007) which show that Japanese and American faces are rated by Mexicans above chance levels. Similar Gendron (2014) shows that Himba participants produce assumed universal patterns of emotion judgment when being cued to emotion concepts, interestingly without concept cues they did not show such universality patterns.

Opposing these results of putative universality Chen (2014) shows that Asian faces were better recognized by Caucasian raters in comparison with Caucasian faces. Likewise Prado and colleagues (2013) show that Caucasian and Chinese Australian raters rated Chinese and Caucasian faces more accurate than Mainland Chinese raters did, most likely due to the adoption of Australian culture. Also Kayyal and Russel (2013a) show that American, English speaking Palestinian and Arabic speaking Palestinian raters only label approximately half of a set of expressed emotions with the same emotion words.

Taking together these results it is obvious that to date no clear conclusion can be drawn about universal and specific cultural aspects of facial emotion expression and recognition, as well as specific encoder-decoder patterns (Scherer et al. 2011). Also recent studies concerning the recognition of facially expressed emotions show inconsistent results concerning this universality-specificity debate.

2.2 **Age differences in facial emotion recognition ability**

Beside cultural aspects that certainly seem to have a certain but yet unclear influence on facial emotion perception (Elfenbein & Ambady, 2002; Scherer et al., 2011) also differences in age seem to be influential regarding the perception and processing of facial stimuli (Kadosh, Johnson, Dick, Kadosh, & Blakemore, 2013; Konar, Bennett, & Sekuler, 2013), the activation of different brain regions (Dundas, Plaut, & Behrmann, 2013; Ebner, Johnson, & Fischer, 2012; Ebner, et al., 2013; Hung, 2014) and the processing of facially expressed emotions (Burianova, Lee, Grady, & Moscovitch, 2013; Horning, 2012; Kellough & Knight, 2012; Lambrecht, Kreifelts, & Wildgruber, 2012; Leime, Neto, Alves, & Torro-Alves, 2013; Lima, Alves, Scott, & Castro, 2014; Lundqvist, Svard, & Fischer, 2013; Ma, Li, Niu, Yu, & Yang, 2013; Mienaltowski, et al., 2013; Noh & Isaacowitz, 2013; Riediger, Voelkle, Ebner, & Lindenberger, 2011; Ruffman, Halberstadt, & Murray, 2009; Ruffman, Henry, Livingstone, & Phillips, 2008; Suzuki & Akiyama, 2013; Sze, 2014; Tottenham, Phuong, Flannery, Gabard-Durnam, & Goff, 2013).

Concerning facial emotion processing age-related differences were also particularly found for the specific process of recognizing facially displayed emotions (Ebner, He, & Johnson, 2011; Ebner, et al., 2012; Ebner, Riediger, & Lindenberger, 2010; Horning, 2012; Isaacowitz & Stanley, 2011; Juen, Huber, & Peham, 2012; Kellough & Knight, 2012; Lambrecht, et al., 2012; Leime, et al., 2013; Ma, et al., 2013; Mienaltowski, et al., 2013; Rhodes & Anastasi, 2012; Riediger, et al., 2011; Ruffman, et al., 2008; Suzuki & Akiyama, 2013; Sze, 2014; West, et al., 2012), for a review see Isaacowitz and Stanley (2011) as well for a meta-analysis see Ruffman and colleagues (2008). In their meta-analysis of 28 data-sets Ruffman and colleagues (2008) show that age seems to have an effect on the recognition of at least some of the basic emotions. Also Weidner (2014) shows that age can be seen as a significant predictor of recognition accuracy. In general studies show that in particular elderly adults and children seem to exhibit poorer recognition rates for facially displayed basic emotions than young adults and middle aged adults (Ebner, et al., 2010; Horning, 2012; Leime, et al., 2013; Suzuki & Akiyama, 2013; Weidner, 2014) and young adults partly showing higher recognition accuracy than middle aged adults (Ebner, et al., 2010) or similar results (Horning, 2012). In general it seems that young adults exhibit best performances in emotion recognition accuracy seemingly due to the fact that elderly adults seem to be more susceptible for the failure to detect facially expressed emotions by instances (Sasson et al., 2010).

Whilst it seems to be obvious that age has an influence on facial emotion perception in general incongruent results can be found regarding specific emotions that seem to be recognized worse or better. Whilst some studies find that facially expressed happiness seems to be uninfluenced by age differences (Ebner, et al., 2010; Ma, et al., 2013; West, et al., 2012) other studies show that also for happiness age differences can be found (Horning, 2012; Suzuki & Akiyama, 2013). Whilst young adults seem to outperform elderly adults in emotion recognition of facially expressed anger (Ebner, et al., 2010; Leime, et al., 2013; Suzuki & Akiyama, 2013), disgust (Ebner, et al., 2010; Suzuki & Akiyama, 2013), sadness (Ebner, et al., 2010; Horning, 2012; Leime, et al., 2013; Suzuki & Akiyama, 2013) and fear (Horning, 2012; Leime, et al., 2013; Suzuki & Akiyama, 2013), also studies can be found that are inconsistent with these results for example showing no such differences for anger (Horning, 2012) or disgust (Horning, 2012). Additionally there seem to be processes besides the explicit emotion recognition and classification that seem to be influenced by age like the recognition of expressions in ambiguous faces (Kellough & Knight, 2012). Furthermore Sasson and

colleagues (2012) show that elderly adults are more likely to tend to label faces with negative emotions.

Taking together these results it seems that age has a certain effect on facial emotion recognition with a tendency of children and elderly adults exhibiting poorer recognition accuracy rates than young and middle aged adults in general. With a closer look however no consistent results regarding combinations of age-groups and specific basic emotions can be found concerning lower recognition accuracy rates.

2.3 **Gender differences in facial emotion recognition capability**

Beside the previously discussed factors culture and age which seem to have specific influence on facial emotion perception and the specific process of emotion recognition it seems that also the gender of a person has to be considered as an additional influential factor on human face perception.

Concerning general facial perception and processing gender differences could be found in healthy persons (Donges, Kersting, & Suslow, 2012; Knyazev, Slobodskoj-Plusnin, & Bocharov, 2010; Rennels & Cummings, 2013; Sommer, Hildebrandt, Kunina-Habenicht, Schacht, & Wilhelm, 2013; Stoesz & Jakobson, 2013; Xu, Yang, Wang, Sun, & Zhao, 2013) as well as in depression or other psychiatric diseases (Bourne & Vladeanu, 2013; Briceno, 2014).

Gender differences could also be found in the processing of facially displayed emotions (Lee, et al., 2013; Torro-Alves, de Oliveira Bezerra, Claudino, & Pereira Cavalcanti, 2013; Wang, 2013) as well as for appendant activation in specific brain regions underlying these processes (Carre, Fisher, Manuck, & Hariri, 2012; Kinoshita, et al., 2012; Pagliaccio, et al., 2013; Rahman & Anchassi, 2012; Schwabe, Hoffken, Tegenthoff, & Wolf, 2013; Tahmasebi, et al., 2012) whilst in other regions no such differences can be found (Derntl, et al., 2012).

Studies show that in general gender differences can be found regarding the specific process of facial emotion recognition (De Carvalho Pinto, Dutra, Filgueiras, Juruna, & Stingel, 2013; Hall, 1978; Hall, Carter, & Horgan, 2000; Hall & Matsumoto, 2004; Hoffmann, Kessler, Eppel, Rukavina, & Traue, 2010; Horning, 2012; Juen, Huber, & Peham, 2012; Knyazev et

al., 2010; Mancini, Agnoli, Baldaro, Bitti, & Surcinelli, 2013; Thayer & Johnson, 2000) whilst other studies couldn't find gender differences regarding facial emotion recognition at all (Hutchison & Gerstein, 2012; Juen, et al., 2012; Melchers, Montag, Markett, & Reuter, 2013; Weidner, 2014).

In their meta analyses Hall (1978) as well as Hall and colleagues (2000) show that females in general produce higher recognition accuracy rates concerning facial emotion recognition than males do. Other studies support these findings showing that female children and adolescents exhibit greater recognition accuracy than male ones (Juen, et al., 2012; Mancini, et al., 2013) and similarly female adults perform more accurate than male adults do (De Carvalho Pinto, et al., 2013; Hall & Matsumoto, 2004; Hoffmann, et al., 2010; Horning, 2012; Thayer & Johnson, 2000). Whilst those studies did find supportive evidence for gender differences in emotion recognition, a number of studies found no or at least partly no hints for a facial emotion recognition advantage for females. Juen and colleagues (2012) showed that whilst gender differences could be found in children and adolescents no differences could be found for preschoolers. Although some studies show no gender differences in adults at all (Calvo & Lundquist, 2008; Hutchison & Gerstein, 2012, Melchers, et al., 2013), other studies show that under specific conditions gender differences could exist. Weidner (2014) couldn't find gender advantages or disadvantages in general although slight benefits for females recognizing neutral faces could be found. Hoffmann and colleagues (2010) showed that highly expressive faces do not trigger gender differences whereas recognition of subtle expressions leads to an advantage for female raters. De Carvalho Pinto and colleagues (2013) show that for the specific emotions anger, fear, surprise and disgust no gender differences could be found whereas for facially expressed sadness women were more accurate than men whilst for happiness men showed higher accuracy rates. Sasson and colleagues (2012) on the other hand indicate a response bias for females regarding the emotions sadness and fear. Other studies concerning facial emotion recognition accuracy didn't show or analyze for gender differences due to methodological reasons like sample compositions including only female subjects (Goeleven, De Raedt, Leyman, & Verschuere, 2008) or unequal distributions of males and females (Langner, et al., 2010; Tottenham, et al., 2009; Tracy, Robins, & Schriber, 2009).

Considering the previously mentioned studies concerning facial emotion recognition and gender differences it becomes obvious that there are a lot of inconsistent results that permit no clear conclusions about gender differences regarding human facial emotion recognition.

2.4 **Databases for emotional face expressions and its characteristics**

Facially expressed emotions have a long tradition and ongoing interest in emotion research (see Adolphs, 2002; Scherer et al., 2011). The fact that there seem to be some emotions having certain facial displays that are cross-culturally recognized at above-chance levels and at the same time inheriting specific cultural variations in expression and recognition (Elfenbein & Ambady, 2002; Scherer et al., 2011), the necessity for standardized and validated emotional face-databases is obvious.

The first widely used database for photographs of static emotional faces was provided by Ekman and Friesen (1976), offering their Pictures of Facial Affect (POFA). This database has been used very often, allowing standardized research on facial emotion perception whilst unfortunately having one big disadvantage, namely its small number of faces (for a discussion see Palermo & Coltheart, 2004). To date there is a growing number of databases providing a set of static photographs displaying facially expressed emotions, for example the Karolinska Directed Emotional Faces (KDEF – Calvo & Lundqvist, 2008; Goeleven, De Raedt, Leyman, & Verschuere, 2008; Lundqvist, Flykt, & Oehman, 1998), Montreal Set of Facial Displays of Emotion (MSFDE - Beaupré, Cheung, & Hess, 2000; Beaupré & Hess, 2005), FACES (Ebner et al., 2010), Japanese and Caucasian Facial Expressions of Emotions (JACFEE – Biehl et al., 1997), Radboud Faces Database (Langner et al., 2010) NimStim set of facial expressions (NimStim – Tottenham et al., 2009) and the University of California, Davis, Set of Emotion Expressions (UCDSEE - Tracy, Robins, & Schriber, 2009).

These mentioned datasets provide information about a number of dimensions based on validation studies such as displayed emotion (obligatory), valence, arousal, intensity, reaction times, attractiveness, clarity, genuineness or reliability, although the emphasis of chosen rating dimensions differs between the databases. The dimensions valence, arousal and dominance (dominance happens to be not typically included in the rating of faces) for the characterization of faces were derived from the approach to rate pictorial material on a number of emotional aspects (Lang, Bradley, & Cuthbert, 1999). In their International Affective Picture System (IAPS) Lang, Bradley and Cuthbert (1999) provide pictorial stimuli with ratings on valence, arousal and dominance. The choice of these dimensions was based on the findings of Osgood, Suci, and Tannenbaum (1957) showing that these three dimensions explain most variance when emotional issues and materials are described. Valence is described as the degree to which someone feels pleasure or discomfort, arousal stands for the

degree of excitement one feels whereas dominance/control means the feeling of being in or out of control concerning emotional stimuli (Lang, Bradley, & Cuthbert, 1999).

Those databases which have been cross-culturally validated show in general middle or high cross-cultural accuracy (Beaupré & Hess, 2005; Biehl et al., 1997, Tottenham et al., 2009) with recognition rates at better-than-chance levels. Looking for validated databases which putatively show high intra-cultural accuracy (Calvo & Lundqvist, 2008; Ebner et al., 2010; Goeleven et al., 2008; Langner et al., 2010; Palermo & Coltheart, 2004; Tracy et al., 2009) valid conclusions can't be drawn due to missing information about nationality and cultural aspects of participants. Some validation studies only give information about participants being students in certain countries (Spanish university - Calvo & Lundqvist, 2008; Belgian university - Goeleven et al., 2008; Dutch university - Langner et al., 2010; Australian university – Palermo & Coltheart, 2004; Tracy et al., 2009) whilst only Ebner and colleagues (2010) report participants being Caucasian native German speakers.

2.5 Theoretical background underlying the choice of stimuli for the Viennese Emotional Stroop Task and implications for their validation

For the construction of the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff 2012) 5 female faces and 5 male faces were randomly drawn from the Karolinska Directed Emotional Faces (KDEF - Lundqvist et al. 1998), each face presenting a straight gaze fearful and straight gaze happy expression. The choice of the photos results from the requirements of the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012) to react to faces displaying either happy or fearful expressions whilst ignoring emotion words that are written straight across the face.

The Karolinska Directed Emotional Faces - KDEF (Lundqvist et al. 1998) contain emotional expressions from 70 Caucasian photo models which have been recorded under standardized conditions (soft, even light; uniform T-shirt colors; fixed image coordinates for eyes and mouths and no salient characteristics like beards, earrings or visible make-up). These emotional faces have been evaluated lately on emotional content (Calvo & Lundqvist, 2008; Goeleven et al., 2008) as well as on intensity and arousal (Goeleven et al., 2008). They show that the recognition accuracy for happy faces is significantly higher (Calvo & Lundqvist, 2008; Goeleven et al., 2008) and for fearful faces significantly lower than for all other facial

displayed emotions (Calvo & Lundqvist, 2008; Goeleven et al., 2008). This goes firm with other findings reporting happy faces being more accurately recognized (Beaupré & Hess, 2005; Biehl et al., 1997; Ebner et al., 2010; Langner et al., 2010; Palermo & Coltheart, 2004; Tottenham et al., 2009) and fearful faces being less accurately recognized than all other emotions (Beaupré & Hess, 2005; Biehl et al., 1997; Palermo & Coltheart, 2004; Tottenham et al., 2009) intra- as well as cross-culturally. Despite the evidence for happy faces eliciting higher recognition rates than fearful faces (Beaupré & Hess, 2005; Biehl et al., 1997 Calvo & Lundqvist, 2008; Ebner et al., 2010; Goeleven et al., 2008; Langner et al., 2010; Palermo & Coltheart, 2004; Tottenham et al., 2009; Tracy et al., 2009) there is a necessity for validating the faces that were randomly chosen for the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012).

For rating the explicit emotion displayed by a face it is common to use multiple choice formats using the generally displayed basic emotions as possible responses (for example Beaupré & Hess, 2005; Calvo & Lundqvist, 2008; Ebner et al., 2010) optionally adding other answers like "?" (Goeleven et al., 2008), "other" (Langner et al., 2010; Palermo & Coltheart, 2004; Tracy et al., 2009), "none of the above/none of these terms is correct" (Tottenham et al., 2009; Tracy et al., 2009 respectively), "no emotion" (Tracy et al., 2009) or additional distractor emotions, for example "excitement" (Tracy et al., 2009), in speeded or unspeeded conditions (Calvo & Lundqvist, 2008). Another possibility for rating the content of a facially expressed emotion is the Facial Action Coding System (FACS; Ekman, Friesen, & Hager, 2002), a technique which is based on the recognition of activation of facial muscles and the intensity of activation. As can be seen in the Method section this technique won't be chosen for the validation of the Viennese Emotion Tasks (Willinger, Schmoeger, Deckert, & Auff, 2012/2013) due to a rather time consuming education process necessary for being able to use FACS and the fact that results could only be generalized onto a population of people who passed such an extensive education.

For rating valence Langner and colleagues (2010) use a 5-point scale (negative to positive), whereas for rating arousal Goeleven and colleagues (2008) use a combination of the Self-Assessment Manikin (SAM; Bradley & Lang, 1994; Lang, 1980) and a 9-point scale (1-"calm" to 9-"aroused"). For rating the intensity of the displayed emotion Likert-scales in different variations are commonly used, for example 5-point (Langner et al., 2010), 7-point (Palermo & Coltheart, 2004), 9-point (Biehl et al., 1997; Goeleven et al., 2008) or continuous

scales (Beaupré & Hess, 2005). Langner and colleagues (2010) also rate their images on attractiveness on a 5-point scale (1-“unattractive” to 5-“attractive”) viewing the neutral expressions of each photo model. Langner and colleagues (2010) argue their additional rating of attractiveness with the fact that studies show that objects are significantly rated better when presented together with straight gaze attractive faces in comparison with less attractive faces (Strick, Holland, & van Knippenberg, 2008).

2.6 Theoretical background underlying the choice of stimuli for the Viennese Affective Flexibility Task and implications for their validation

For the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013) 2 male and 3 female models were randomly chosen from KDEF (Lundqvist et al. 1998) each one displaying a happy, angry, sad and fearful straight gaze expression. This was done due to the fact that conducting the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013) participants have to match faces according to their expressed emotions without being cued to the explicit emotion categories happiness, anger, sadness or fear.

Regarding the KDEF faces Calvo and Lundqvist (2008) show that for happy faces the recognition accuracy is significantly higher than for angry, sad and fearful faces, whilst for sad and fearful faces the recognition rate is significantly lower. Goeleven and colleagues (2008) show similar results with the difference of fear being the only emotion being rated significantly less accurate than anger and sadness. In general however validation studies on emotional databases show inconsistent results regarding differences in recognition accuracy with the only constant result of happiness being rated most accurately (Beaupré & Hess, 2005; Biehl et al., 1997; Ebner et al., 2010; Langner et al., 2010; Palermo & Coltheart, 2004; Tottenham et al., 2009, Tracy et al., 2009).

Due to the previously mentioned reasons of universality and specificity on facial emotion perception as well as inconsistent encoder-decoder patterns (Elfenbein & Ambady, 2002; Elfenbein et al., 2007; Scherer et al., 2011) the necessity of a proper validation of the Viennese Affective Flexibility Task items (Willinger, Schmoeger, Deckert, & Auff, 2013) is implicated.

2.7 **The necessity for validating the Viennese Emotional Stroop Task and Viennese Affective Flexibility items**

As previously discussed three factors can be found most potentially influencing the perception and processing of emotional faces as well as the specific process of recognition of facially expressed emotions, namely culture, gender and age.

The faces taken from the Karolinska Directed Emotional Faces - KDEF (Lundqvist et al. 1998) have to date been evaluated by two validation studies (Calvo & Lundqvist, 2008; Goeleven et al., 2008). Whilst they provide ratings of emotion recognition accuracy and other dimensions typically rated with faces they lack to provide sample-specific information about cultural or sub-cultural aspects, native language as well as nationality, therefore it is not possible to draw conclusions about encoder-decoder universalities or specificities. As discussed before there is an unclear body of acquired knowledge concerning universal as well as cross-cultural, sub-cultural, territorial, literate or preliterate aspects and inconsistent encoder-decoder patterns regarding facial emotion perception and evaluation (Elfenbein & Ambady, 2002; Elfenbein et al., 2007; Scherer et al., 2011) which definitely implies the validation of the KDEF faces on a representative Austrian and in Austria living German speaking sample. Although for gender Calvo and Lundqvist (2008) find no differences in face ratings at all the previously discussed findings show that gender can't be ignored as a possible factor confounding emotional face processing despite of the inconsistence in results. Whilst Calvo and Lundqvist (2008) as well as Goeleven and colleagues (2008) didn't analyze for differences concerning the age of raters a number of studies show that age should be considered a possible influential factor on the processing and recognition of facially expressed emotions.

Summarizing these results the necessity for a proper evaluation of the chosen material for the Viennese Emotion Tasks (Willinger, Schmoeger, Deckert, & Auff, 2012/2013) on a representative sample comprising Austrians and in Austria living German speaking persons with an additional focus on age- and gender differences is implied.

3. The perception and processing of language, emotional words and emotional properties of words

In research of human perception and information processing the usage of verbal stimuli is widely distributed due to a lot of advantages they bear. One of the greatest advantages of using words for experimental stimuli is the possibility to control a series of quantifiable factors like e.g. number of letters, number of syllables, number of phonemes, number of orthographic neighbors, frequency of usage (Soares, Comesana, Pinheiro, Simoes, & Frade, 2012; Voe, Jacobs, & Conrad, 2006; Voe, et al., 2009) and to avoid potentially influential picture properties like color, contrast, brightness (Forsythe, Mulhern, & Sawey, 2008) and other factors which increase visual complexity, in order to provide greater experimental control (Soares et al., 2012).

Also in emotion research verbal material is widely used, using emotion words or words with different emotional properties to assess emotional information processing in different paradigms like the emotional go/no-go Task (for example Chiu, Holmes, & Pizzagalli, 2008; Gole et al., 2012) or the emotional stroop Task (Egner, Etkin, Gale, & Hirsch, 2008; Etkin et al., 2006; Etkin, Prater, Hoeft, Menon, & Schatzberg, 2010; Mathews & MacLeod, 1985; McKenna, 1986). Besides the previously mentioned quantifiable linguistic properties words also bear the potential to be used as experimental stimuli whilst heeding other psycholinguistic factors according to which words can be classified or controlled for. Likewise to pictorial (Lang, Bradley, & Cuthbert, 1999) and auditory material (Bradley & Lang, 1999a) words can be characterized according to different emotional dimensions like valence, arousal or dominance (ANEW - Bradley & Lang, 1999) or other dimensions like e.g. imagery or familiarity which aren't natural properties of these words but must be evaluated in validation studies. In the following section a number of word-databases and their attendant validation studies are presented showing which dimensions are typically collected when creating such databases and how they are collected.

3.1 **Databases for words and its characteristics**

Bradley and Lang (1999) created and validated the Affective Norms for English Words (ANEW), a database providing normative emotional ratings for 1034 words (nouns, verbs and adjectives) in English language. Bradley and Lang (1999) based their validation on the theoretical framework of Osgood, Suci and Tannenbaum (1957) who showed that most variance in emotional assessment can be explained by three dimensions namely valence (judgment on pleasantness), arousal (degree of activation or excitement) and dominance (degree of control a subject feels regarding a stimulus). In order to do so they used the semantic differential method. On basis of these results Bradley and Lang (1999) conducted a validation study rating the chosen words on dimensions valence, arousal and dominance using combinations of 9-point scales and the Self-Assessment Manikin (SAM; Bradley & Lang, 1994; Lang, 1980). For each dimension specific SAMs with different visual aspects were used. The Affective Norms for English Words (ANEW) can be accounted as the most used and most often adapted word database providing normative ratings for valence, arousal and dominance. ANEW has been translated into Spanish (Redondo, Fraga, Padron, & Comesna, 2007) and Portuguese (Soares et al., 2012), rated by discrete emotional categories (Stevenson, Mikels, & James, 2007) and extended (Warriner, Kuperman, & Brysbaert, 2013).

In the style of ANEW a number of databases have been developed in different languages, for example English (Cortese & Fugett, 2004; Eilola & Havelka, 2010; Stevenson, Mikels, & James, 2007; Warriner, Kuperman, & Brysbaert, 2013), Spanish (Ferré, Guasch, Moldovan, & Sanchez-Casas, 2012; Redondo et al., 2007), French (Gilet, Gruehn, Studer, & Labouvie-Vief, 2012), Finnish (Eilola & Havelka, 2010; Soederholm, Hayry, Laine, & Karrasch, 2013), German (Briesemeister, Kuchinke, & Jacobs, 2011; Gruhn & Smith, 2008; Kanske & Kotz, 2010; Lahl, Goeritz, Pietrowsky, & Rosenberg, 2009; Voe, Jacobs, & Conrad, 2006), Italian (Barca, Burani, & Arduino, 2002) Portuguese (Soares et al., 2012) and Dutch (Moors et al., 2013). Most of these databases validate the chosen words on arousal, valence and dominance, although other rating dimensions like imagery/imageability (Cortese & Fugett, 2004; Gilet et al., 2012; Gruhn & Smith, 2008; Voe et al., 2006), concreteness (Eilola & Havelka, 2010; Ferré et al., 2012; Kanske & Kotz, 2010; Lahl et al., 2009), age of acquisition (Barca, Burani, & Arduino, 2002; Moors et al., 2013; Soares et al., 2012), familiarity (Eilola & Havelka, 2010; Ferré et al., 2012), offensiveness (Eilola & Havelka, 2010), or discrete categories (Briesemeister, Kuchinke, & Jacobs, 2011; Stevenson, Mikels, & James, 2007) are often used

instead or additionally. For dominance seemingly being a dimension explaining less variance than arousal and valence (Bradley & Lang, 2000; Russell, 2003) a number of validation studies conduct ratings for the latter two excluding ratings of dominance (Eilola & Havelka, 2010; Ferré et al., 2012; Gilet et al., 2012; Kanske & Kotz, 2010; Lahl et al., 2009; Soederholm et al., 2013; Voe, Jacobs, & Conrad, 2006). Those studies which include ratings of dominance show positive correlations between dominance and valence indicating that positive words are rated as being more controllable and/or vice versa (Soares et al., 2012; Warriner, Kuperman, & Brysbaert, 2013).

In 2009, Voe and colleagues created the Berlin Affective Word List Reloaded (BAWL-R), a database with more than 2000 words in German language, that contains a number of psycholinguistic indexes, for example number of letters, number of syllables, frequency (per million words), but also ratings of emotional valence, emotional arousal and imageability of words (Voe, et al., 2009). In their validation study (Voe, et al., 2009) they rated the chosen words on emotional valence using a 7-point scale with a total range between -3.0 (extremely negative) and +3.0 (extremely positive), on imageability using a 7-point scale ranging between 1 (low imageability) and 7 (high imageability) whereas for the rating of emotional arousal they chose a 5-point scale with a total range from 1 (low arousal) to 5 (high arousal). For the rating of emotional arousal they used a combination of the previously mentioned 5-point scale and the Self-Assessment Manikin (SAM; Bradley & Lang, 1994; Lang, 1980) due to the fact that the German word for arousal [“Erregung”] potentially implies sexual associations (Voe, et al., 2009). Using the SAM (Bradley & Lang, 1994) they argue that these connotations can be attenuated or avoided (Voe, et al., 2009). The BAWL-R has been validated on a sample of 200 psychology students at the Freie Universität Berlin and the Katholische-Universität Eichstätt-Ingolstadt, the majority most likely being native German speakers and most likely of German nationality although these and other information about cultural aspects can’t be found in the article.

3.2 **Cultural differences and the processing of words' emotional properties**

Whilst translating ANEW into other languages (Redondo et al., 2007; Soares et al., 2012) both validation studies implied ratings with representative cultural (Redondo et al., 2007; Soares et al., 2012) and explicit language-specific samples (Soares et al., 2012). Comparing these translation studies with the original ANEW validation study similar results and result

patterns can be found between American, Spanish and Portuguese samples although some cultural differences regarding the ratings of valence, arousal and dominance seem to exist (Bradley & Lang, 1999; Redondo et al., 2007; Soares et al., 2012). Eilola and Havelka (2010) collected affective norms for Finnish and British nouns within samples of native Finnish speakers and British English speakers showing that most of the ratings (valence, arousal, offensiveness, concreteness and familiarity) were strongly correlated between the two languages, similar cross-cultural patterns are also shown by Soederholm and colleagues (2013).

These results are partly in line with other studies showing cross-cultural differences (Ishii, Reyes, & Kitayama, 2003; Marosi, Rivas, Yanez, & Bernal, 2002; Russell & Sato, 1995; Triandis & Osgood, 1958; Wierzbicka, 2010) and differences between languages (Pavlenko, 2008; Wierzbicka, 2010) regarding emotional connotations of words and processing of these as well as cultural differences in the description of explicit emotion categories (Goddard, 1997; Kayyal & Russell, 2013b; Panayiotou, 2002; Rusch, 2004; Russell & Sato, 1995; Schmidt-Atzert & Park, 1999). Whilst these studies show cross-cultural and language-specific differences regarding emotion words and emotional connotations of words other studies show no such differences concerning emotion category labeling (Hupka, Lenton, & Hutchison, 1999) or emotion word encoding cross-culturally (Moore, Romney, Hsia, & Rusch, 1999) and in mono- and bilinguals (Rusch, 1997). For an overview about universality and cultural specificity of emotion processing see Elfenbein & Ambady, 2002.

Although these studies show similarities between rather close as well as rather distinct cultures regarding emotional ratings of words, they also show that these ratings can at the same time been seen as culture sensitive and dependent on nationality or even territorial belonging. These results lead to the necessity of validating chosen verbal stimulus material on representative samples. Gole and colleagues (2012) for example searched for differences in emotional perception and information processing between high worry and low worry groups applying an emotional go/no-go task using words from the BAWL (Voe et al., 2006). Participants were recruited via announcement on an Austrian and/or German University, leading to a sample most likely consisting of native German speakers with Austrian and/or German nationality, although these are only assumptions. Following the main task participants had to rate the words on valence and arousal using the Self-Assessment Manikin (Bradley & Lang, 1994). Results showed that within this specific sample aversive words were

rated as more negative and more arousing than neutral words. This goes in line with studies indicating that ratings of valence and arousal show a typical U-shaped relationship with negative and positive words having higher arousal ratings than neutral words and negative words having the highest arousal ratings in the majority of studies (Eilola & Havelka, 2010; Ferré et al., 2012; Kanske & Kotz, 2010; Kanske & Kotz, 2011; Moors et al., 2013; Redondo et al., 2007; Soares et al., 2012; Soederholm et al., 2013; Voe, et al., 2009).

3.3 **The influence of age on words' emotional connotations**

In the previous chapters of this diploma thesis studies were presented which hinted that beside cultural aspects also age seems to be in some way influential on facial emotion recognition.

Likewise to culture it seems to be that age also has a certain impact on the processing of emotional verbal stimuli (Abbassi, 2012; Ashley & Swick, 2009; Castro & James, 2014; Dunajska, Szymanik, & Trempala, 2012; Kensinger, 2008; Leclerc & Kensinger, 2011; Mammarella, Borella, Carretti, Leonardi, & Fairfield, 2013; Molnar et al., 2013; Murray & Kensinger, 2013; Silk et al., 2009; Thapar & Rouder, 2009; Thomas & Hasher, 2006; Wurm, Labouvie-Vief, Aycock, Rebucal, & Koch, 2004; Yang & Hasher, 2011) as well as on brain activation underlying such processes (Leclerc & Kensinger, 2011; Molnar et al., 2013). In particular age seems to influence the processing of emotional words in the course of language production (Castro & James, 2014), memory (Kensinger, 2008; Leclerc & Kensinger, 2011; Mammarella et al., 2013; Murray & Kensinger, 2013; Thapar & Rouder, 2009; Thomas & Hasher, 2006; Yang & Hasher, 2011) and reaction time in different paradigms (Ashley & Swick, 2009; Dunajska et al., 2012; Silk et al., 2009; Wurm et al., 2004). Also concerning the usage of emotional words age differences could be found (Marosi, Rivas, Yanez, & Bernal, 2002).

Among the previously mentioned standardized word databases only one could be found which deliberately looked for age differences concerning the rating of words according to the already mentioned dimensions (Gilet et al., 2012). Gilet and colleagues (2012) looked for similarities as well as for differences in emotional word ratings between young aged, middle aged and elderly adults and showed that in general similar rating patterns could be found. Despite these similarities they also showed slight differences between the age groups concerning valence rating with middle aged adults tending to rate words more positively than

younger and elderly adults. With a closer look no differences between the age groups concerning mean valence rating could be found. Also for the rating dimensions arousal and imageability differences could be found with young adults rating words less arousing than middle aged and elderly adults as well as elderly adults rating words as more imaginable than middle aged adults which themselves rated higher than the young ones did. Elderly adults furthermore showed higher correlations between valence and arousal ratings than the younger groups did.

Also Gruhn and Smith (2008) show that elderly adults tend to higher positive ratings than younger adults do, at the same time rating positive words more arousing, less controllable and more easily imaginable. They also show that elderly adults rate by trend negative words as less arousing and more controllable. Syssau and Monnier (2009) show that valence ratings of words change in the course of childhood showing with age groups of 5-,7-, and 9-year olds that with increasing age the number of positive word ratings decreases, whilst exhibiting a stable percentage of negative ratings they show an increasing ratio of neutral words.

Summarizing these results age differences seem to exist for rating words according to their emotional properties as well as other descriptive psycholinguistic dimensions although it seems that these differences range in smaller magnitudes.

3.4 **Gender-specific influences on word ratings**

Besides cultural and age-related aspects which both seem to influence emotional word processing albeit in unclear ways and seemingly different magnitudes also gender could most likely be an influential factor.

Studies show gender differences in different emotional word processing procedures (Breitberg et al., 2013; Fussell, Rowe, & Mohr, 2012; Garcia-Garcia, Dominguez-Borras, SanMiguel, & Escera, 2008; Glenberg, Webster, Mouilso, Havas, & Lindeman, 2009), e.g. attentional biases (Breitberg et al., 2013) or reaction times (Dunajska, Szymanik, & Trempala, 2012; Garcia-Garcia et al., 2008; Gohier et al., 2013) with additional differences in attendant physiological activity (Bellace, Williams, Mohamed, & Faro, 2012; Coney & Fitzgerald, 2000; Hofer et al., 2007; Hyona & Haikio, 2005; Landis, 2006). Interestingly also hormones were found to play a certain role in gender differences regarding emotional word processing (Breitberg et al.,

2013). Also concerning the usage of emotional words (Dewaele & Pavlenko, 2002; Marosi, Rivas, Yanez, & Bernal, 2002) differences between males and females could be found.

Only a handful of articles focus on gender differences regarding emotional word ratings (Bauer & Altarriba, 2008; Soares et al., 2012; Syssau & Monnier, 2009), putting their attention mainly on ratings of valence. Syssau & Monnier (2009) for example show that within age-groups of 5-, 7-, and 9-year old children no gender differences concerning valence ratings could be found. With adults on the other hand Bauer and Altarriba (2008) show that females rate concrete words significantly higher on the dimension emotionality in comparison with male adults. Due to the fact that the term emotionality, ranging between 'not at all emotional' and 'very emotional', isn't described exactly it is rather difficult to draw clear conclusions.

From the previously mentioned word databases only one validation study reports slight gender differences (Soares et al., 2012) showing that males and females differ in their arousal ratings with women showing higher ratings. For the rating of valence Soares and colleagues (2012) find no gender differences at all.

Likewise to the influential impact of age also for gender only unclear results can be found at the most indicating the possibility of only slight gender differences.

3.5 Theoretical background underlying the choice of stimuli for the Viennese Emotional Go/No-go Task and implications for their validation

For the construction of the Viennese Emotional Go/No-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) 100 aversive and 100 neutral nouns were chosen from the Berlin Affective Word List Reloaded (BAWL-R) (Voe, et al., 2009). The selection was based on word length and the emotional valence of the words, first selecting words with 5, 6 or 7 letters. For the neutral stimuli the possible item pool was reduced to those words, which emotional valence were rated around the midpoint of the 7-point scale, exactly within a range of -0.1 and +0.1. From the resulting words, 100 were randomly chosen. For the aversive stimuli an item pool was created containing words with emotional valence ratings within the most negative 25 per cent of the scale exactly within a range of -3.0 and -1.6. Within this range all words with ratings between -3.0 and -2.0 were automatically chosen (56 words),

from the remaining words (valence ratings ranging between -2.0 to -1.6) 44 were chosen randomly. Thus 100 neutral words with 5, 6 or 7 letters and an emotional valence between -0.1 and +0.1 and 100 negative words with 5, 6 or 7 letters and an emotional valence between -3.0 and -1.6 were chosen for the Viennese Emotional Go/No-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012). On basis of the dimensional ratings provided by the BAWL-R (Voe, et al., 2009) statistical analysis were conducted (Willinger, Schmoeger, Deckert, & Auff, 2012). T-tests (independent samples) show significant differences between neutral and negative words regarding emotional valence and arousal scores. For emotional valence neutral words ($M = 0.005$, $SD = 0.06$) were significantly rated less negatively than negative ($M = -2.09$, $SD = 0.33$) words ($t(198) = 63.27$, $p \leq 0.0001$) as for arousal neutral words ($M = 2.55$, $SD = 0.571$) were significantly rated less arousing than negative ($M = 3.88$, $SD = 0.43$) words ($t(198) = -18.67$, $p \leq 0.0001$).

Thus the sample of 200 words that were drawn out of the 2000 words in the BAWL-R (Voe et al., 2009) in order to be used in the Viennese Emotional Go/No-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) statistically show the presumed emotional properties with predetermined negative words showing more negativity in valence rating and higher ratings in arousal than the predetermined neutral words. These analysis were conducted using the ratings provided by the validation study of Voe and colleagues (2009) showing that the assumed properties of the words can be determined for the previously described sample Voe and colleagues (2009) used. As discussed in the previous sections the processing of emotional words is most likely influenced by the factors culture, age and gender. With Voe and colleagues (2009) presenting no clear characteristics of the sample concerning native language, nationality and other cultural aspects no meaningful inferences can be drawn for the usage of the BAWL-R words in samples comprising other cultural characteristics like e.g. a sample of Austrians or German speaking people living in Austria. Also having discussed that age and gender seem to be influential on the processing of emotional words Voe and colleagues (2009) miss to analyze for age- or gender differences concerning the ratings of the BAWL-R words.

In order to use the chosen words from the BAWL-R as stimuli for the Viennese Emotional Go/No-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) as well as for further usage in other future tasks the necessity for evaluating the chosen words on a representative Austrian

and German speaking in Austria living sample and analyzing the results additionally for age- and gender differences in implied.

Empirical work

4. Method

4.1 **Procedure**

First participants were informed about the nature and course of the study. Subsequently they gave written informed consent to participate in this study which had been approved by the Ethical Commission of the Medical University of Vienna.

Following participants completed a demographic data form containing questions about a number of population parameters which are more or less likely to potentially influence the perception and processing of emotional verbal material as well as facially expressed emotions. The demographic data form contained questions about age, gender, place of birth, country/district/city in which participants spent most of their childhood/adolescence/adulthood, native language, additional acquired languages, frequencies of usage of these languages, handedness, religion, education, profession, illnesses, and medicaments currently taking.

In order to screen for anxiety or depression symptoms participants subsequently completed the German version of the Hospital anxiety and depression scale – HADS-D (Herrmann-Lingen, Buss, & Snaith, 2011) which has been shown to assess caseness and symptom severity of anxiety disorders and depression quite well, also in the general population (Bjelland, Dahl, Haug, & Neckelmann, 2002). Furthermore participants were also asked to complete the German version of the Difficulties in emotion regulation scale – DERS (Gole, Koechel, Schaefer, & Schienle, 2012; Gratz & Roemer, 2004) so as to screen for problems in emotion regulation.

After completing these questionnaires participants were asked to rate a number of words as well as subsequently a number of human faces in two separate computer tasks. First participants were asked to conduct Computer Task I (programmed using E-Prime® 2.0) in

order to rate the words chosen for the Viennese Emotional Go/No-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) on a number of dimensions. Subsequently participants were asked to conduct Computer task II (programmed using E-Prime® 2.0) so as to rate the pictures of facially expressed emotions chosen for the Viennese Emotional Stroop and Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013). Computer Task I as well as Computer Task II will be described in the following sections.

Participation in this validation study approximately took 2 – 2 ½ hours whilst participants did not get any financial compensation.

4.2 **Subjects**

90 healthy participants took part in this validation study and were recruited personally or via notices. Exclusion criteria were neurological or psychiatric diseases as well as deficient German language skills. All participants had normal or corrected-to-normal vision.

In order to analyze for age differences as well as for gender differences the sample was divided into three age groups comprising young aged adults (20-30 years of age), middle aged adults (40-50 years) and elderly adults (60-70 years). The composition of the three age groups was determined according to age-related differences regarding cognition (Brickman et al., 2006; Grady, Springer, Hongwanishkul, McIntosh, & Winocur, 2006; Hildebrandt, Wilhelm, Schmiedek, Herzmann, & Sommer, 2011), regional white matter volume (Brickman et al., 2006), perception and processing regarding facial information (Hildebrandt et al., 2011) as well as perception of facially expressed emotions (Ebner et al., 2010). The exact age-span inside the groups as well as the age differences between the groups were set more stringently than in the previously mentioned literature in order to allow more meaningful results concerning similarities and/or differences between those age groups as well as to potentially improve the homogeneity inside each group.

4.2.1 Age

The age of the whole sample ranged between 20 and 70 years (M = 45.2; SD =16.7) (see figure 1 and table 1).

The young adult group had a mean age of 24.7 years (SD = 2.4) and consisted of 15 men (Mean age = 24.8, SD = 2.2) and 15 women (Mean age = 24.5, SD = 2.6). The middle aged adult group had a mean age of 46.3 years comprising 15 men (Mean age = 46.6, SD = 3.1) and 15 women (Mean age = 46, SD = 3.2). Elderly adults had a mean age of 64.7 years (SD = 3.6) and similarly consisted of 15 men (Mean age = 65.9, SD = 2.6) and 15 women (Mean age = 63.5, SD = 4.1). For an overview see table 1).

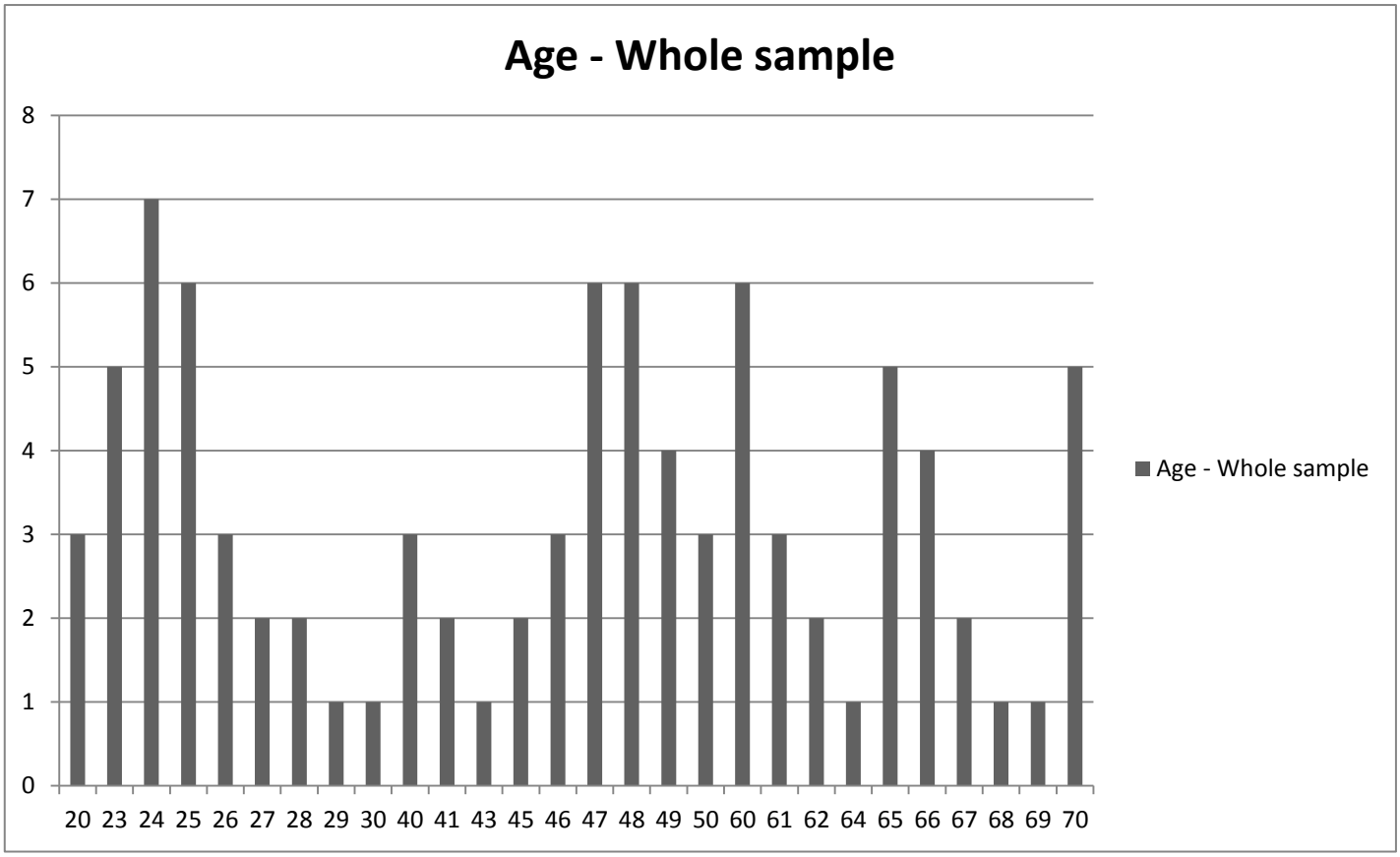


Figure 1: Age distribution in the sample.

Table 1: Mean age (M) and standard deviations (SD) are presented for the whole sample, for the young adult group (Young), middle aged adult group (Middle) and elderly adults (Elderly) as well as for males and females and the gender-age-subgroups.

Gender	Age Group		Age
Male	Young	M	24.8
		SD	2.2
	Middle	M	46.6
		SD	3.1
	Elderly	M	65.9
		SD	2.6
	Overall	M	45.8
		SD	17.2
Female	Young	M	24.5
		SD	2.6
	Middle	M	46
		SD	3.2
	Elderly	M	63.5
		SD	4.1
	Overall	M	44.7
		SD	16.4
Overall	Young	M	24.7
		SD	2.4
	Middle	M	46.3
		SD	31.1
	Elderly	M	64.7
		SD	3.6
	Overall	M	45.2
		SD	16.7

4.2.2 Gender

The whole sample (N=90) consisted of 45 females (50%) and 45 males (50%) (see figure 2 and table 2).

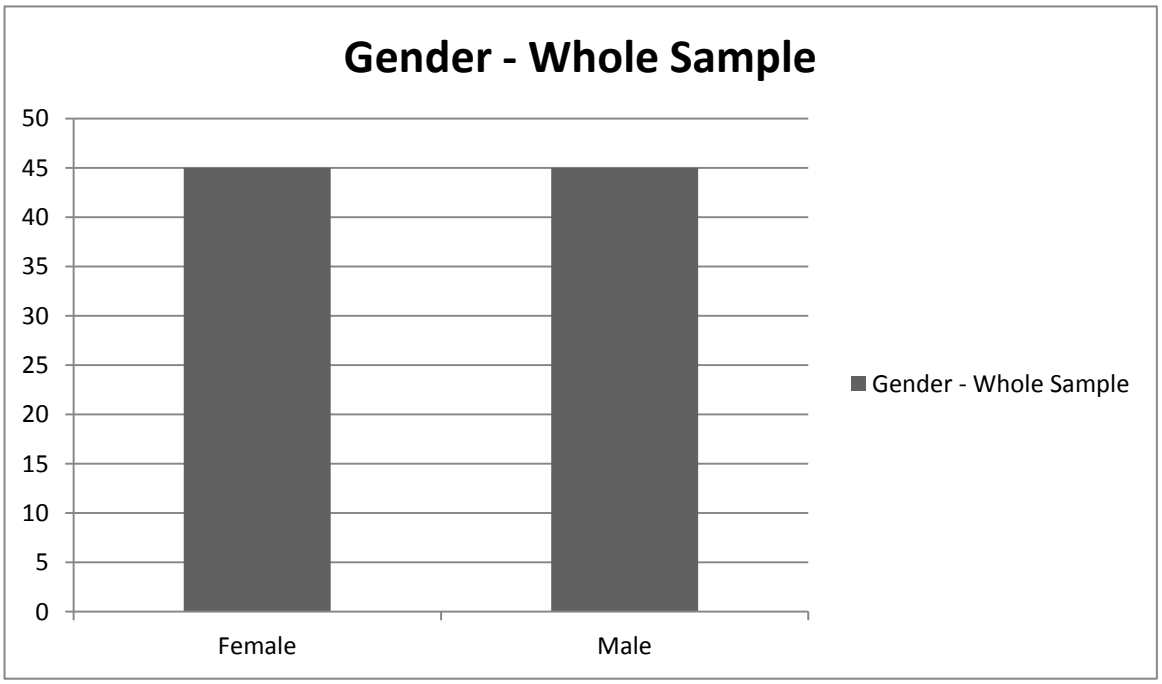


Figure 2: Gender distribution in the whole sample.

The uniform distribution of gender across age groups was analyzed using a χ^2-test. The sample showed an uniform distribution of gender ($\chi^2(2) = 0$, $p = 1$).

Table 2: Cross-tabulation for age groups and gender

Age Group	Gender		Overall
	Male	Female	
Young aged adults (20-30)	15	15	30
Middle aged adults (40-50)	15	15	30
Eldery aged adults (60-70)	15	15	30
Gesamt	45	45	90

4.2.3 Land of birth and country in which most time of the adulthood was spent

84 participants (93%) were born and raised in Austria whilst the other six participants were born in Germany (2), Russia, Hungary, Iceland and Australia (see figure 3). All 84 native Austrians spent nearly their whole life living in Austria whilst of the six non-native Austrians all spent at least the great majority of their adulthood living in Austria (two of them being raised in Austria since childhood) (see figure 4).

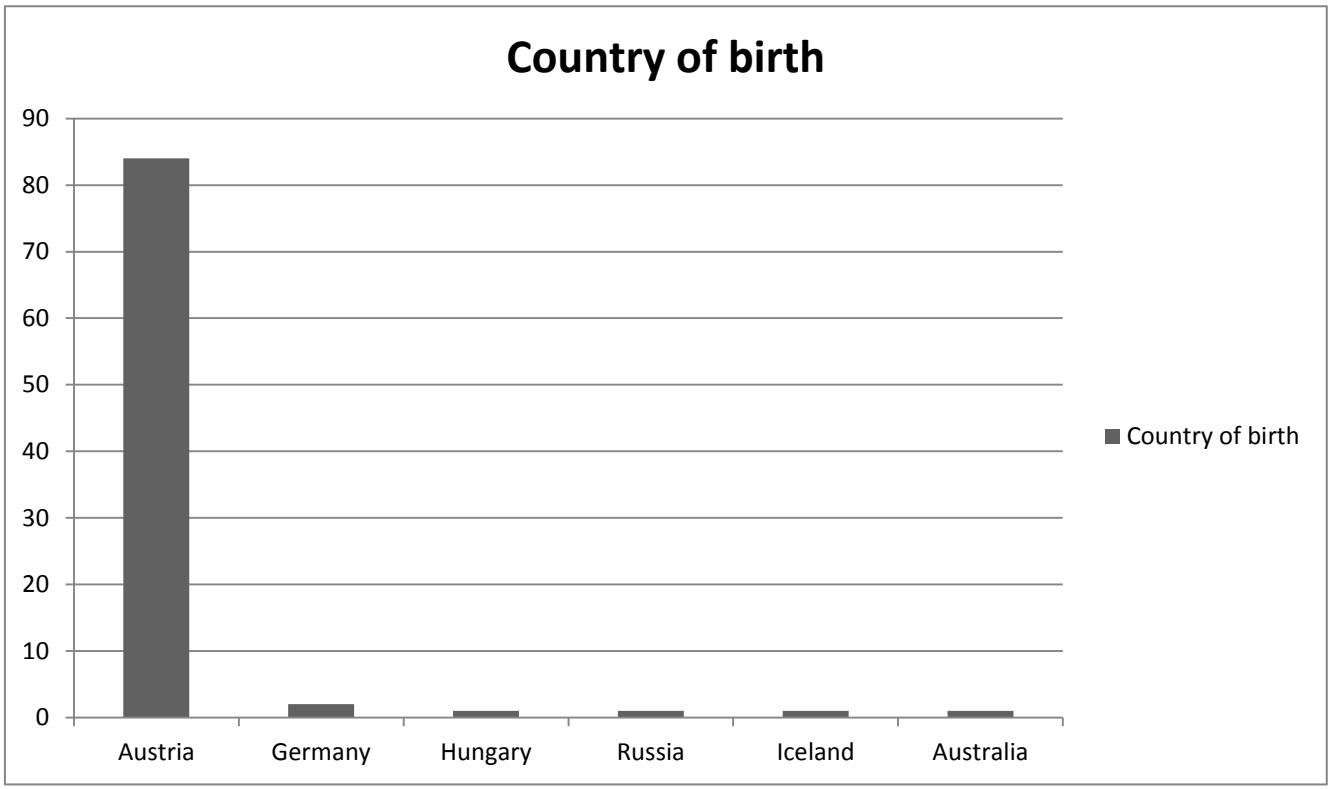


Figure 3: Places of birth of the participants.

The uniform distribution of countries of birth across age groups was analyzed using a χ^2-test. The sample did show an uniform distribution of countries of birth across age groups ($\chi^2(10) = \leq .9{,}07$, $p = .525$).

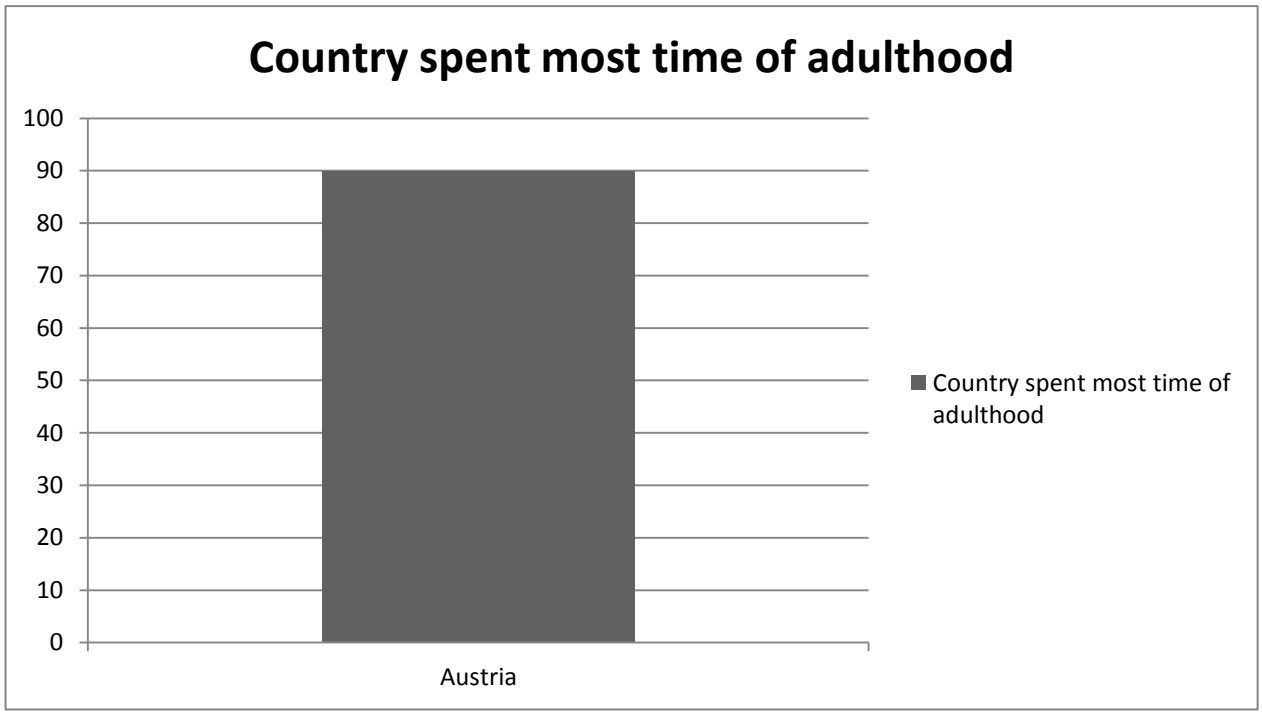


Figure 4: Country in which participants spent most of their adulthood.

4.2.4 Native language

All 90 participants had good or very good German language skills for German being either their native language (97%) or having acquired it as second language using it very often (3%) (see figure 5).

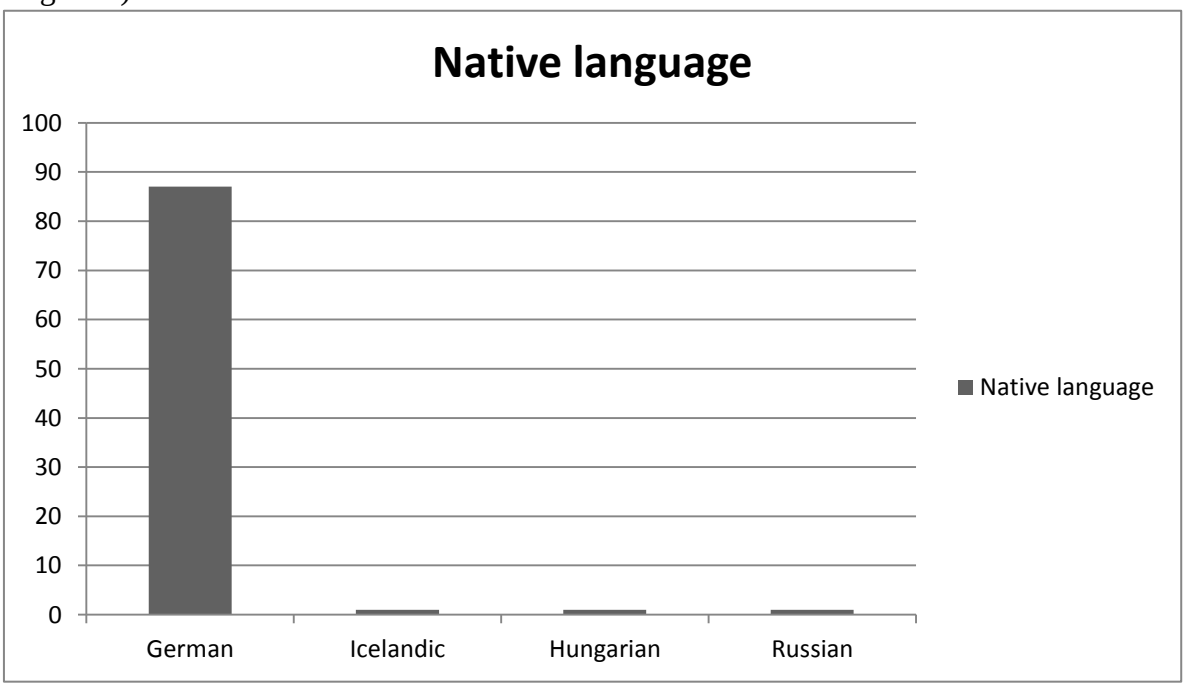


Figure 5: Native language of all participants.

4.2.5 Occupation

55 Participants (61%) were employed at the time point of study participation whilst 35 participants (39%) did not have a job (anymore) (see figure 6).

Figure 6: Current employment situation of all participants.

The uniform distribution of occupation status across age groups was analyzed using a χ^2-test. The sample did not show an uniform distribution of occupation status ($\chi^2(2) = \leq 43.85$, $p \leq .0001$) (see table 3).

Table 3: Cross-tabulation for age groups and occupation status

Age Group	Occupation		Overall
	No	Yes	
Young aged adults (20-30)	6	24	30
Middle aged adults (40-50)	3	27	30
Eldery aged adults (60-70)	26	4	30
Gesamt	35	55	90

4.2.6 Education

14 participants declared ‘Pflichtschule’ being their highest degree of education (15%), three declared ‘Lehre’ (3%), 19 participants had finished a ‘Fachschule’ (22%), one participant had passed the ‘Studienberechtigungsprüfung’ (1%), 32 had declared ‘Matura/Abitur’ being their highest degree of education (36%), eight finished an ‘Akademie (9%), three declared ‘Fachhochschule’ (3%) and 10 had a university degree (11%) (see figure 7).

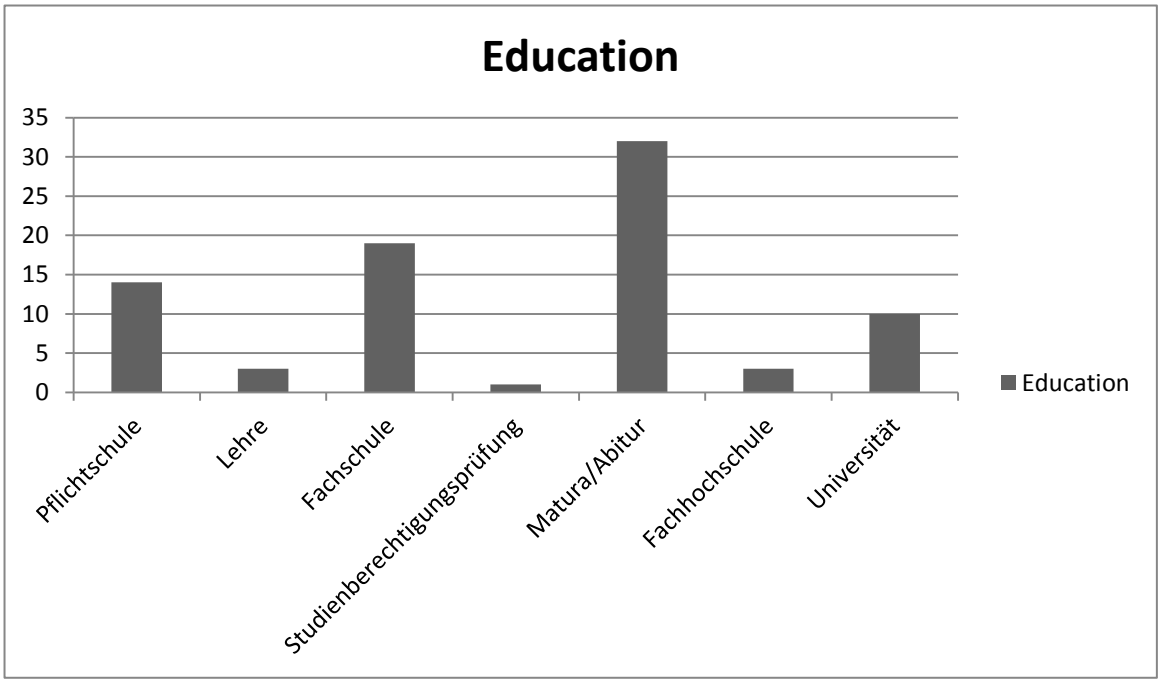


Figure 7: Education degrees within the sample.

The uniform distribution of education across age groups was analyzed using a χ^2-test. The sample did not show an uniform distribution of education ($\chi^2(16) = \leq 40.285$, $p = .001$) (see table 4).

Table 4: Cross-tabulation for age groups (Young, Middle, and Elderly – adults) and education.

Age Group	Education								
	Pflicht-Schule	Lehre	Fachschule	Studien-berechtigungs-prüfung	Matura/ Abitur	Akademie	Fachhochschule	Universität	Overall
Young	2	0	2	1	17	0	0	8	30
Middle	7	1	7	0	8	3	3	1	30
Eldery	5	2	10	0	7	5	0	1	30
Gesamt	14	3	19	1	32	8	3	10	90

4.2.7 DERS- and HADS-D-scores

Conducting the German version of the Difficulties in emotion regulation scale – DERS (Gole, Koechel, Schaefer, & Schienle, 2012; Gratz & Roemer, 2004) offering a scoring range between 36 and 180 participants had a mean score of 68 (SD = 15) which seems to be a normal result for healthy persons compared with other studies (Gole, Koechel, Schaefer, & Schienle, 2012; Gratz & Gunderson, 2006). In the German version of the Hospital anxiety and depression scale – HADS-D (Herrmann-Lingen, Buss, & Snaith, 2011) participants had a mean score of 2.6 (SD = 2.1) for the depression scale and a mean score of 4.5 (SD = 2.7) for the anxiety scale with three participants showing slightly increased scores.

4.3 **Materials**

The following two computer tasks were programmed by the author of this diploma thesis using E-Prime® 2.0.

4.3.1 Computer Task I - Validation of the Viennese Emotional Go/No-go Task items

In the first of two consecutive Computer Tasks participants were shown the words chosen for the Viennese Emotional Go/No-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) one after another with negative and neutral words being randomized in order. The target word was always shown on the upper half of the screen being accentuated by a larger font size than the sequently shown questions. Although being presented in random order participants had to rate each word they read according to the dimensions valence, arousal, dominance, imagery and familiarity in this fixed order.

Participants were additionally instructed that if a word had several and/or ambiguous meanings they could choose the meaning of the word that would be predominant for them personally. This was done due to the aim of obtaining authentic word ratings on valence and the other dimensions by a sample of Austrians and German speaking people living in Austria. Each dimension and its appendant rating scale according to which the words were rated will be described in the following sections.

a) *Valence rating*

First participants had to rate the shown word on valence. Seeing the word on the upper half of the screen the question "*How negative (unpleasant) or positive (pleasant) are your feelings reading this word?*" was presented directly beneath. Similar to other studies concerning emotional word ratings (Bradley & Lang, 1999; Ferré et al., 2012; Gole et al, 2012; Kanske & Kotz, 2010; Redondo et al., 2007; Soares et al, 2012) a combination of the Self-Assessment Manikin (SAM - Bradley & Lang, 1994) and a 9-point Likert scale with valence ranging from -4 ('very negative') to +4 ('very positive') with 0 ('neutral') as midpoint was used. The visualizing SAMs showed analogous emotional mimics. 'Very negative' (-4) had to be chosen when the target word triggered feelings like e.g. being very unhappy, annoyed, despaired, desperate, anxious, hopeless or bored. 'Very positive' (+4) had to be chosen when the target word triggered feelings like e.g. being very happy, amused, pleased or hopeful. The specific bipolar modality was chosen in line with Voe and colleagues (2009).

b) *Arousal rating*

After rating a word on valence participants had to rate the same target word on the dimension arousal. Whilst still viewing the word on the upper half of the screen the question: "*How great is your inner arousal reading this word?*" was to read below. For the rating itself a combination of the Self-Assessment Manikin (SAM - Bradley & Lang, 1994) and a 9-point Likert scale with arousal ranging from 1 ('low arousal') to 9 ('high arousal') with 5 ('middle') as midpoint was presented beneath the question as it is usual in other studies rating emotional words for arousal (Bradley & Lang, 1999; Ferré et al., 2012; Gole et al., 2012; Kanske & Kotz, 2010; Redondo et al., 2007; Soares et al., 2012). The SAMs visualized representative shades of arousal by varying mimic and physical arousal display. 'Low arousal' (1) had to be chosen when the target word triggered states of feeling e.g. very relaxed, calm, unagitated, sluggish or bored. 'High arousal' (9) had to be chosen when the target word triggered states of e.g. strong excitement, stimulation, nervousness, agitation, huffishness, tension or stress. Like in the validation study of Voe et al. (2009) a unipolar scale modality was chosen for the rating of arousal.

c) *Dominance rating*

After rating the word for valence and arousal participants were asked to rate the given word for dominance. In the style of Bradley & Lang (1999) participants had to rate for the degree of

feeling controlled or feeling being in control (dominant) regarding the target word. Still seeing the previously rated word on the upper half of the screen the question appeared below: "*How controlled or dominant do you feel reading this word?*". Like in other studies (Bradley & Lang, 1999; Redondo et al., 2007; Soares et al., 2012) a combination of the Self-Assessment Manikin (SAM - Bradley & Lang, 1994) and a 9-point Likert scale with dominance ranging from 1 ('controlled') to 9 ('dominant') with 5 ('balanced') as midpoint was used. The size of the SAMs varied analogously to the degree how much one feels controlled (small SAM) or dominant (large SAM). 'Controlled' (1) had to be chosen when the target word triggered feelings e.g. of strongly being controlled, swayed, intimidated, being made submissive but also being shepherded or lead. 'Dominant' (9) had to be chosen when the target word triggered feelings of e.g. being very decisive, controlling, influential, important or autonomous.

d) *Imagery rating*

Following the ratings for valence, arousal, and dominance participants rated the word on the dimension imagery. As in other studies (Cortese & Fugett, 2004; Gilet et al., 2012; Voe et al., 2006) the rating to which degree the word can be imagined was conducted using a pointed scale. Although usually a 7-point scale is used for imagery ratings in this study a 9-point scale ranging from 1 ('very difficult to imagine') to 9 ('very easy to imagine') with 5 ('middle') as midpoint was used because of the idea of commonly using 9-point scales for the complete rating process. Whilst participants still saw the word on the upper half of the screen the question "*How easily does this word evoke a visual, acoustic or other mental image?*" was presented underneath. 'Very difficult to imagine' (1) had to be chosen when participants had great difficulties with either mentally seeing, hearing, smelling, tasting or touching what they associated with the target word whereas 'very easy to imagine' (9) had to be chosen when participants had no difficulties at all mentally imagining the target word.

e) *Familiarity rating*

In the last step participants had to rate the word on the dimension familiarity. In order to do so they still saw the previously rated word with the question below: "*How often does this word occur in everyday written or spoken language?*". In the style of Eilola & Havelka (2010) as well as Ferré et al. (2012) a 9-point scale with 1 ('very seldom') till 9 ('very often') with 5 ('middle') as midpoint was used. Participants had to choose 'very seldom' (1) when they had the feeling that the target word never or only very seldom occurs in their everyday language

either being heard, spoken or written. ‘Very often’ (9) had to be chosen when they had the feeling that the target word very often occurs in their everyday language. When participants asked what everyday language was exactly meant to be they were told that this rating should be based on their own personal experience with this word.

All word ratings were conducted using Computer Task I (programmed using E-Prime® 2.0) as participants had to select and click on a position on or between the rating scale points of which they thought it would adequately represent the shown word on each dimension. Participants had also the possibility of directly clicking on the corresponding SAMs or the spaces between them. For an overview of the rating scales of Computer Task I see the appendix (A2-A6).

4.3.2 Computer Task II - Validation of the Viennese Emotional Stroop Task and Viennese Affective Flexibility Task items

In Computer Task II participants first had to rate the attractiveness of the photo models that were chosen for the Viennese Emotional Stroop and the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012/2013) whilst looking at the faces displaying neutral expressions. Afterwards they were asked to identify the different emotions shown by the models as well as rate them on the dimensions intensity, arousal, and dominance. Subsequently they had to conduct a second emotion identification task. All of these validation operations were done viewing photo models displaying a straight gaze. This order of ratings has partly been adapted from Langner and colleagues (2010).

a) *Attractiveness rating*

Similar to the validation study of Langner and colleagues (2010) validating their Radboud Faces Database, participants first had to rate the 15 chosen photo models (5 females and 5 males for the Viennese Emotional Stroop Task as well 3 females and 2 males for the Viennese Affective Flexibility Task) displaying a neutral expression on the dimension attractiveness. Participants saw each face on top of the screen with the question below: “*How attractive do you think this person is?*”. For the attractiveness rating a 9-point scale ranging from -4 (‘unattractive’) to +4 (‘very attractive’) with 0 (‘middle’) as midpoint was used, presented right beneath the face and the question. In contrast to Langner and colleagues (2010) which

used a 5-point scale, a 9-point scale was used in this study due to the simple idea of commonly using 9-point scales for all rating dimensions.

b) *Emotion recognition task I*

After rating the attractiveness of the 15 photo models participants were instructed how to rate each of the following faces, each one displaying a certain emotion.

First participants had to identify the displayed emotion of a face viewing it on top of the screen with the question right below: "*Which emotion is displayed?*". Participants had to identify an either fearful, happy, angry, sad or neutral face using a multiple choice format given the answer options 'happiness' [Freude], 'fear' [Angst], 'anger' [Ärger], 'sadness' [Trauer], 'neutral' [Neutral] and 'none of the mentioned' [Keine der Genannten]. The answer options 'happiness', 'fear', 'anger' and 'sadness', each one most likely representing a so called 'basic' emotion, were chosen due to the fact that these should be the emotions which were meant to be displayed by the photo models. Conducting the Viennese Emotion tasks (Willinger, Schmoeger, Deckert, & Auff, 2012/2013) each of these emotions has to be recognized. Therefore as answer options only the emotions were chosen, for which they could be mistaken in the course of conducting the tasks additionally adding the answer option 'none of the mentioned'.

Rating the chosen photo models also faces with neutral expressions were displayed, provided with the same answer options likewise to the rating of emotional faces. This was done so as to analyze for rating patterns which could potentially show the labeling of neutral faces with certain emotions.

The answer "None of the mentioned" was chosen in the style of Tracy and colleagues (2009) because of the methodological critics, limitations and response artifacts of forced choice methods (Frank & Stennet, 2001) in order to enable participants only to label a face when they are of the opinion that an existing emotion is adequately expressed and the right answer can be chosen (Tottenham et al., 2009). The answer options were presented in a row right beneath the question, whilst the exact positions of the answer options were randomized from face to face. Participants had to click on the answer option of which they thought it would be the right answer for the facially displayed expression.

Instantaneously after picking one answer in the emotion identification task participants had to rate the very same face for the following dimensions: intensity, arousal, and dominance.

c) *Intensity rating*

Right after choosing an answer in the emotion recognition task participants were asked to rate the intensity of the displayed emotion. In order to do so they were shown the same face they just had rated on top of the screen with the question "*How intense is this emotion displayed?*" right underneath. In the style of Biehl and colleagues (1997) as well as Goeleven and colleagues (2008) 9-point scale ranging from 0 ('no intensity') till 8 ('high intensity') with 4 ('middle') as midpoint was used for the rating of intensity and was presented directly below the question. 'No intensity' (0) had to be chosen when participants had the feeling that no emotion was displayed at all whereas 'high intensity' (8) had to be chosen when they had the feeling that the emotion was displayed with a high degree of intensity. Participants had to rate the intensity of the displayed emotion independently of the answer they previously chose in the emotion recognition task in order to guarantee a proper intensity judgment even if the displayed emotion was before identified as "None of the mentioned".

d) *Arousal rating*

Right after identifying the displayed emotion and rating it for the dimension intensity participants had to conduct a rating of their personal inner arousal provoked by seeing this facially displayed emotion. They still saw the face on top of the screen whilst now the question "*How great is your inner arousal viewing this emotion?*" was presented right below. In the style of Goeleven and colleagues (2008) a combination of the Self-Assessment Manikin (SAM; Bradley & Lang, 1994; Lang, 1980) and a 9-point scale ranging from 1 ('low arousal') till 9 ('high arousal') with 5 ('middle') as midpoint was used. Analogous to the 9-point scale the SAMs visualized different graduations of arousal. For a detailed description of the scale points 1 ('low arousal') and 9 ('high arousal') see section 4.3.1.b.

e) *Dominance rating*

Following the identification task as well as the rating of intensity and arousal participants still saw the same face which now they had to rate for dominance (feeling of being controlled by this stimulus or being in control of it) in the last step. Being one of the dimensions potentially explaining a certain amount of variance in emotion perception (Bradley & Lang 1999; Osgood, Suci and Tannenbaum, 1957) dominance was included in the validation of the

Viennese Emotional Stroop and the Viennese Affective Flexibility Task items (Willinger, Schmoeger, Deckert, & Auff, 2012/2013). Beneath the face the question "*How controlled or dominant do you feel viewing this emotion?*" was presented. In the style of Bradley and Lang (1999) participants rated for dominance using a combination of the Self-Assessment Manikin (SAM; Bradley & Lang, 1994; Lang, 1980) and a 9-point scale ranging between 1 ('controlled') and 9 ('dominant') with 5 ('balanced') as midpoint placed right below the question. Visualizing the different degrees of feeling being controlled (small SAM) or being in control (large SAM) the SAMs differed in size and mimic. For a detailed description of the scale points 1 ("controlled") and 9 ("dominant") see section 4.3.1.c.

The ratings of attractiveness, intensity, arousal and dominance in Computer Task II were conducted clicking on a position on or between the rating scale points (or SAMs) for which participants thought that it would adequately represent the displayed emotion on the corresponding dimension. Given 15 photo models displaying neutral expressions on the attractiveness rating as well as displaying 5 different expressions on the recognition task and on the intensity, arousal and dominance ratings participants had to conduct 315 trials (no time limit) so far in Computer Task II.

f) *Emotion recognition task II*

The successful conduction of the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012) requires amongst others the ability to identify facially expressed emotions as either fearful or happy whilst ignoring an emotion word that is written across the face. Therefore it is necessary to rate these photo models in an additional discrimination task (no time limit) providing the two answer options "fear" and "happiness" per expressed emotion.

In order to additionally validate the Viennese Emotional Stroop Task items (Willinger, Schmoeger, Deckert, & Auff, 2012) participants had to conduct 'Emotion recognition task II' (emotion discrimination) after successfully conducting the other tasks and ratings of Computer Task II. This second emotion recognition task required the capability to identify the displayed emotions shown by the previously described 15 photo models as either happy or fearful expressions. The faces were shown on top of the screen whilst below the question was presented: "*Which emotion is displayed?*". The answer options 'fear' [Angst] and 'happiness' [Freude] were placed right below the question. Each face-emotion combination was presented

twice whilst the answer options were counterbalanced in left-right position and inverted for the second presentation of each face. In the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012) there are no direct repetitions of the same photo model expressing the same emotion (for example twice in a row model number 1 expressing fear) in order to avoid repetition priming effects (Mayr, Awh, & Laurey, 2003; see also Egner, Etkin, Gale, & Hirsch, 2008; Etkin, Egner, Peraza, Kandel, & Hirsch, 2006; Etkin, Prater, Hoeft, Menon, & Schatzberg, 2010). Due to this fact there was no direct repetition of the same face expressing the same emotion in this discrimination Task.

This second emotion recognition task was conducted clicking on the chosen answer options for each face. Given 15 photo models each presenting 2 emotion expressions twice participants had to conduct 60 discriminations in 'Emotion recognition task II'.

For an overview of the rating scales as well as figures of the multiple choice task see the appendix (A2-A9).

5. Hypothesis and statistics

5.1 **Hypotheses for Computer Task I - Validation of the Viennese Emotional Go/No-go Task items**

H1(1): There is a significant difference in the valence rating of neutral and negative words.

H1(2): There is a significant difference in the valence rating of neutral and negative words within age and gender groups.

H1(3): There is a significant difference in the arousal rating of neutral and negative words.

H1(4): There is a significant difference in the arousal rating of neutral and negative words within age and gender groups.

H1(5): There is a significant correlation between the ratings of valence and arousal for neutral words.

H1(6): There is a significant correlation between the ratings of valence and arousal for negative words.

H1(7): There is a significant correlation between the ratings of valence and dominance for neutral words.

H1(8): There is a significant correlation between the ratings of valence and dominance for negative words.

For hypotheses *H1(1)* and *H1(3)* paired t-tests will be conducted whereas for hypotheses *H1(2)* and *H1(4)* two factorial MANOVAs with age group and gender as independent variables will be applied. For hypotheses *H1(5)*, *H1(6)*, *H1(7)* and *H1(8)* Pearson correlations will be calculated. Results with a p-value <.05 will be considered as statistically significant.

5.2 Hypotheses for Computer Task II - Validation of the Viennese Emotional Stroop Task and Viennese Affective Flexibility Task items

H1(9): For all emotions as well as for neutral faces in the emotion recognition task (emotion recognition I) the recognition rate will be higher than the chance level (6 answer options, resulting in a chance level 1/6 = 0.16).

H1(10): For emotions happiness and fear in the emotion discrimination task (emotion recognition II) the recognition rate will be higher than the chance level (2 answer options, resulting in a chance level 1/2 = 0.5).

H1(11): There is a significant difference in overall facial emotion recognition accuracy rates regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(12): There is a significant difference in recognition accuracy concerning emotions anger, fear, happy and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(13): There is a significant difference in overall intensity rating of facially expressed emotions anger, fear, happiness and sadness (no neutral faces included) regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(14): There are significant differences in intensity ratings of facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(15): There is a significant difference in overall arousal rating of facially expressed emotions anger, fear, happiness and sadness (no neutral faces included) regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(16): There are significant differences in arousal ratings of facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(17): There is a significant difference in overall dominance rating of facially expressed emotions anger, fear, happiness and sadness (no neutral faces included) regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(18): There is a significant difference in dominance rating of facially expressed emotions regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(19): There is a no significant difference in recognition of facially expressed fear or happiness in the discrimination task (emotion recognition II) regarding age groups (young, middle and elderly adults) as well as gender (male, female).

H1(20): There are significant correlations between the dimensions intensity, arousal and dominance concerning the facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces.

H1(21): There is a significant correlation between the ratings of attractiveness and intensity of facially expressed emotions anger, fear, happiness and sadness as well as neutral faces.

H1(22): There is a significant correlation between the ratings of attractiveness and arousal of facially expressed emotions anger, fear, happiness and sadness as well as neutral faces.

H1(23): There is a significant correlation between the ratings of attractiveness and dominance of facially expressed emotions anger, fear, happiness and sadness as well as neutral faces.

For hypotheses *H1(1)* and *H1(2)* frequencies of chosen answer options will be calculated. For hypotheses *H1(3)*, *H1(5)*, *H1(7)*, *H1(9)* two factorial ANOVAs with age group and gender as independent variables will be conducted. In order to analyze hypotheses *H1(4)*, *H1(6)*, *H1(8)*, *H1(10)* and *H1(11)* two factorials MANOVAs with age group and gender as independent variables will be applied. Hypotheses *H1(12)*, *H1(13)*, *H1(14)* and *H1(15)* will be analyzed conducting Pearson correlations. Results with a p-value <.05 will be considered as statistically significant.

6. Results

6.1 Results for Computer Task I - Validation of the Viennese Emotional Go/No-go Task items

H1(1): **There is a significant difference in the valence rating of neutral and negative words.**

A paired t-test was conducted for analyzing differences in valence rating of negative and neutral words.

There was a significant difference between negative and neutral words in valence rating ($t(89 = -28.104$, $p \leq .0001$) with neutral words ($M = 515.11$, $SD = 33.69$) being rated significantly higher in valence than negative words ($M = 270.77$, $SD = 74.37$) (see table 5).

H1(2): **There is a significant difference in the valence rating of neutral and negative words within age and gender groups**.

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and valence ratings of negative and positive words as dependent variables was conducted.

There was a significant main effect for gender on the valence rating of negative words ($F(1,84) = 4.188$, $p = .044$) whilst there was no significant main effect for gender for neutral words ($F(1,84) = .052$, $p = .82$). Furthermore there was no significant main effect for age on the valence rating of negative ($F(2,84) = .09$, $p = .914$) and neutral words ($F(2,84) = 2.504$, $p = .088$) and no significant interaction effect of gender and age on the valence rating of negative ($F(2,84) = 1.249$, $p = .292$) and neutral words ($F(2,84) = 1.253$, $p = .291$).

Males ($M = 286.64$, $SD = 81.46$) rated negative words significantly more positive than females ($M = 253.89$, $SD = 63.53$), for an overview see table 5.

H1(3): **There is a significant difference in the arousal rating of neutral and negative words.**

A paired t-test was calculated for analyzing differences in arousal rating of negative and neutral words.

There was a significant difference between negative and neutral words in arousal rating ($t(89 = 15.277$, $p \leq .0001$) with negative words ($M = 498.08$, $SD = 170.33$) being rated significantly higher in arousal than neutral words ($M = 275.72$, $SD = 117.39$) (see table5).

H1(4): **There is a significant difference in the arousal rating of neutral and negative words within age and gender groups.**

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and arousal ratings of negative and positive words as dependent variables was conducted.

There were no significant main effects for age concerning the arousal rating of negative words ($F(2,84) = .304$, $p = .739$) and neutral words ($F(2,84) = .263$, $p = .769$). Moreover there were no significant main effects for gender regarding the arousal rating of negative words ($F(2,84) = .2.887$, $p = .093$) and neutral words ($F(2,84) = 1.368$, $p = .245$) as well as no significant

interaction effects on the arousal rating of negative ($F(2,84) = .384$, $p = .682$) and neutral words ($F(2,84) = .469$, $p = .627$), for an overview see table 5.

H1(5): **There is a significant correlation between the ratings of valence and arousal for neutral words.**

A Pearson correlation was conducted with valence and arousal ratings of neutral words.

There was no significant coefficient of determinations between valence and arousal ratings for neutral words ($R^2 = .004$, $r = -.064$, $p = .552$).

H1(6): **There is a significant correlation between the ratings of valence and arousal for negative words.**

A Pearson correlation was calculated with valence and arousal ratings of negative words.

There was a significant coefficient of determinations between valence and arousal ratings for negative words ($R^2 = .3$, $r = -.548$, $p \leq .0001$).

Table 5: Means (M) and standard deviations (SD) for valence ratings and arousal ratings for negative words as well as neutral words are presented for the young adult group (Young), middle aged adult group (Middle) and elderly adults (Elderly) as well as for males and females and the gender-age-subgroups.

Gender	Age Group		Valence Rating for Negative Words	Valence Rating for Neutral Words	Arousal Rating for Negative Words	Arousal Rating for Neutral Words
Male	Young	M	296,07	512	439	300,27
		SD	67,91	19,52	123,44	104,83
	Middle	M	299,13	517,6	472,07	298,6
		SD	64,24	35,6	130,59	109,28
	Elderly	M	264,73	513,3	491,27	272,27
		SD	106,49	28,33	204,24	133,32
	Overall	M	286,64	514,31	467,44	290,38
		SD	81,46	28,01	154,99	114,54
Female	Young	M	249,6	499,67	521,33	276,67
		SD	65,91	22,97	185,90	122,83
	Middle	M	247,53	532,4	556,93	237,13
		SD	51,74	46,45	146,57	98,43
	Elderly	M	267,53	515,67	507,87	269,4
		SD	73,52	38,59	213,24	138,83
	Overall	M	254,89	515,911	528,71	261,07
		SD	63,53	38,87	180,94	119,66
Overall	Young	M	272,83	505,83	480,17	288,47
		SD	69,87	21,86	160,6	112,84
	Middle	M	273,3	525	514,5	267,87
		SD	63,03	41,33	143,06	106,86
	Elderly	M	266,13	514,5	499,57	270,83
		SD	89,92	33,29	205,33	133,74
	Overall	M	270,77	515,11	498,08	275,72
		SD	74,37	33,69	170,33	117,39

H1(7): **There is a significant correlation between the ratings of valence and dominance for neutral words.**

A Pearson correlation was conducted with valence and dominance ratings of neutral words.

There was no significant coefficient of determinations between valence and dominance ratings for neutral words ($R^2 = .001$, $r = -.038$, $p = .720$).

H1(8): **There is a significant correlation between the ratings of valence and dominance for negative words.**

A Pearson correlation was calculated with valence and dominance ratings of negative words.

There was a significant coefficient of determinations between valence and dominance ratings for negative words (R^2 = .05, r = .229, p = .03).

6.2 Results for Computer Task II - Validation of the Viennese Emotional Stroop Task and Viennese Affective Flexibility Task items

H1(9): **For all emotions as well as for neutral faces in the emotion recognition task (emotion recognition I) the recognition rate will be higher than the chance level (6 answer options, resulting in a chance level 1/6 = 0.16).**

For analyzing the recognition accuracy percentages for each emotion as well as for neutral faces frequencies of each answer option will be calculated.

Regarding the accuracy percentages all emotions namely happiness (89%), anger (84%), sadness (60%) and fear (52%) as well as neutral faces (73%) were recognized above the chance level (see table 6).

Recognizing facially expressed happiness following answer options were chosen: happiness (91.3%), neutral (3.9%), none of the mentioned (3.1%), sadness (1.0%), fear (.4%) and anger (.2%) (see figure 8).

Concerning the recognition of facially expressed anger the answers anger (83.7%), none of the mentioned (10.1%), sadness (3.9%), fear (1.3%), neutral (.9%) and happiness (.1%) were chosen (see figure 9).

In the course of recognizing facially expressed sadness following answer options were chosen: sadness (74.9%), none of the mentioned (9.3%), neutral (6.3%), anger (6.0%), fear (3.3%) and happiness (.1%) (see figure 10).

Recognizing fearful faces the answer options fear (56.9%), none of the mentioned (27.1%), sadness (8.5%), anger (4.0%), neutral (2.7%) and happiness (.7%) were chosen (see figure 11).

Regarding the recognition of neutral faces the following answers were chosen: neutral (85%), sadness (5.3%), anger (4.2%), none of the mentioned (4.0%), happiness (1.1%) and fear (.4%) (see figure 12).

For an overview see table 6 as well figure 13.

H1(10): **For emotions happiness and fear in the emotion discrimination task (emotion recognition II) the recognition rate will be higher than the chance level (2 answer options, resulting in a chance level 1/2 = 0.5).**

For analyzing the recognition accuracy percentages for each facially expressed fear and happiness frequencies of each answer option (fear or neutral) will be calculated.

Regarding the accuracy percentage fear (92%) as well as happiness (93%) were recognized above the chance level.

Regarding the recognition of facially expressed fear the answer option fear (94%) was chosen more often than happiness (6%).

Recognizing facially expressed happiness the answer option happiness (98.4%) was chosen more often than fear (1.6%), for an overview see table 6 and figure 14.

Table 6: *Upper half of the table*: Overall emotion recognition accuracy for each emotion as well as chosen answer options per emotion in Emotion Recognition task I (Computer Task II) presented in percentage for the whole sample. *Lower half of the table*: Overall emotion recognition accuracy for emotions fear and happiness as well as chosen answer options per emotion in Emotion Recognition task II (Computer Task II) presented in percentage for the whole sample.

Task	Facial Expression	Overall Emotion Recognition Accuracy in Percent	Percentage of Answer Option *Anger*	Percentage of Answer Option *Fear*	Percentage of Answer Option *Happiness*	Percentage of Answer Option *Neutral*	Percentage of Answer Option *Sadness*	Percentage of Answer Option *None of the mentioned*
Emotion Recognition I	Anger	84	83.7	1.3	.1	.9	3.9	10.1
	Fear	52	4	56.9	.7	2.7	8.5	27.1
	Happiness	89	.2	.4	91.3	3.9	1	3.1
	Neutral	73	4.2	.4	1.1	85	5.3	4
	Sadness	60	6	3.3	.1	6.3	74.9	9.3
Emotion Recognition II	Fear	92		94	6			
	Happiness	93		1.6	98.4			

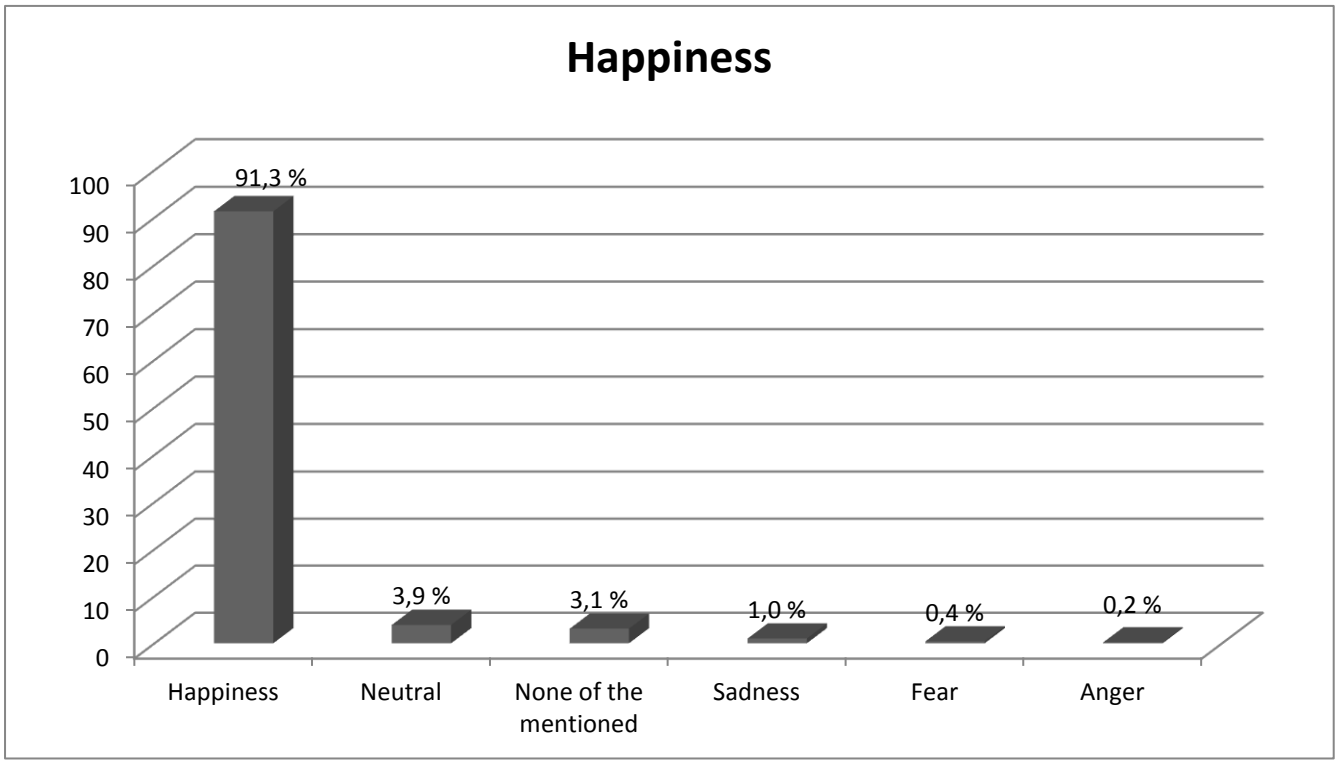


Figure 8: Chosen answer options for facially expressed happiness in Emotion Recognition I, presented in percentages for the whole sample.

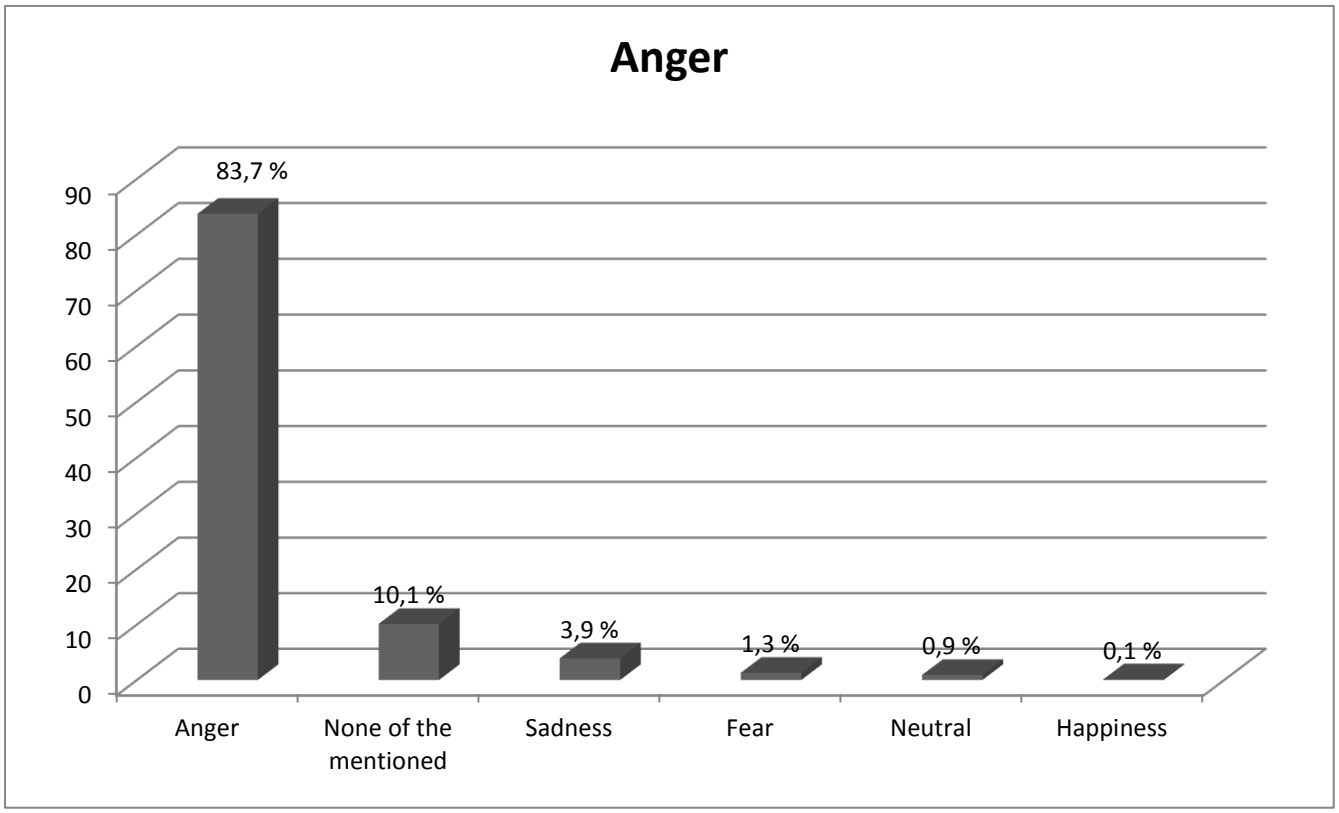


Figure 9: Chosen answer options for facially expressed anger in Emotion Recognition I, presented in percentages for the whole sample.

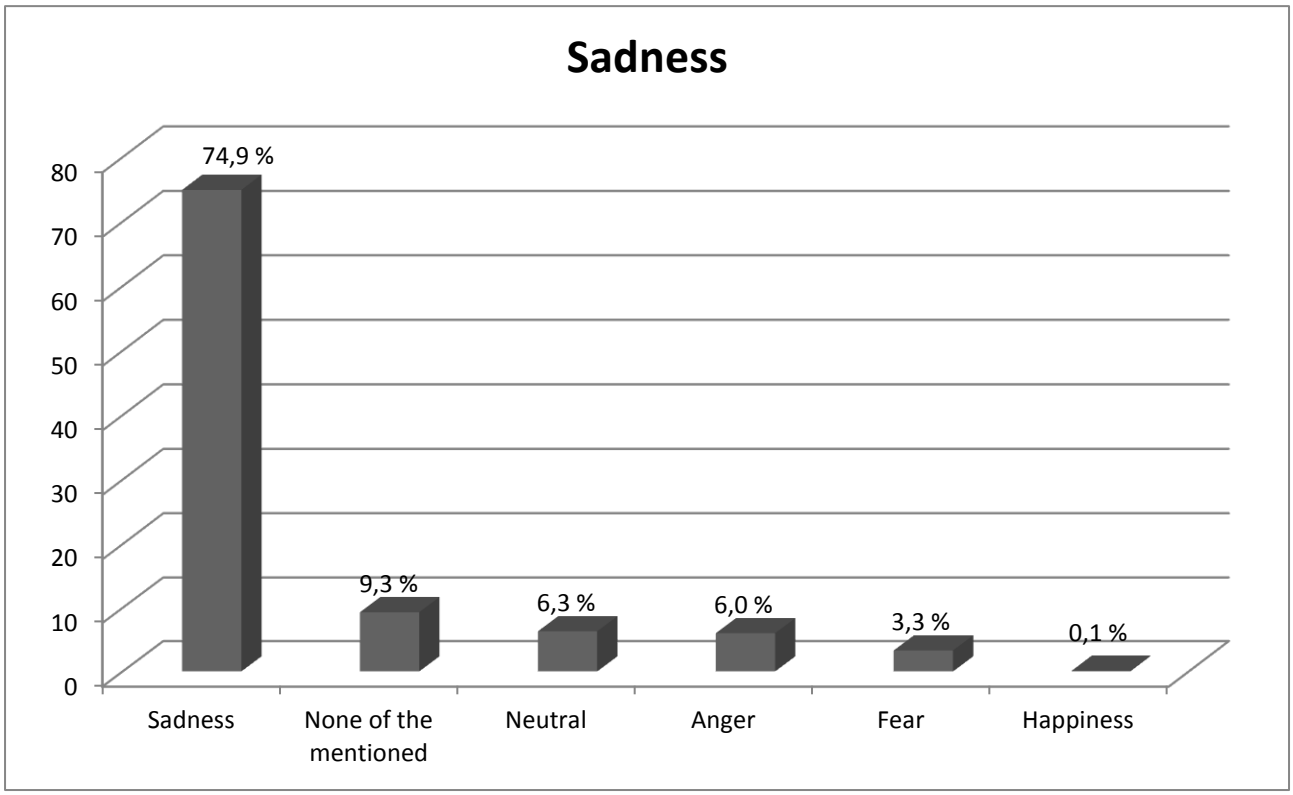


Figure 10: Chosen answer options for facially expressed sadness in Emotion Recognition I, presented in percentages for the whole sample.

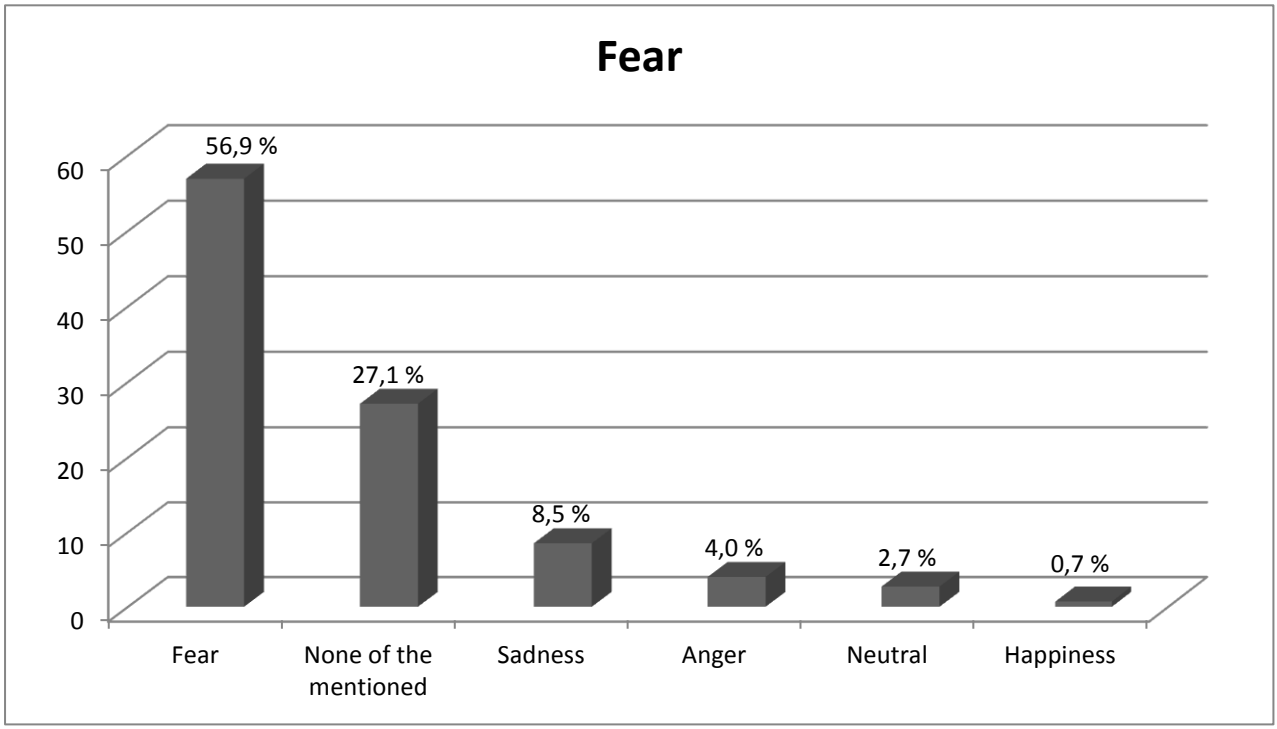


Figure 11: Chosen answer options for facially expressed fear in Emotion Recognition I, presented in percentages for the whole sample.

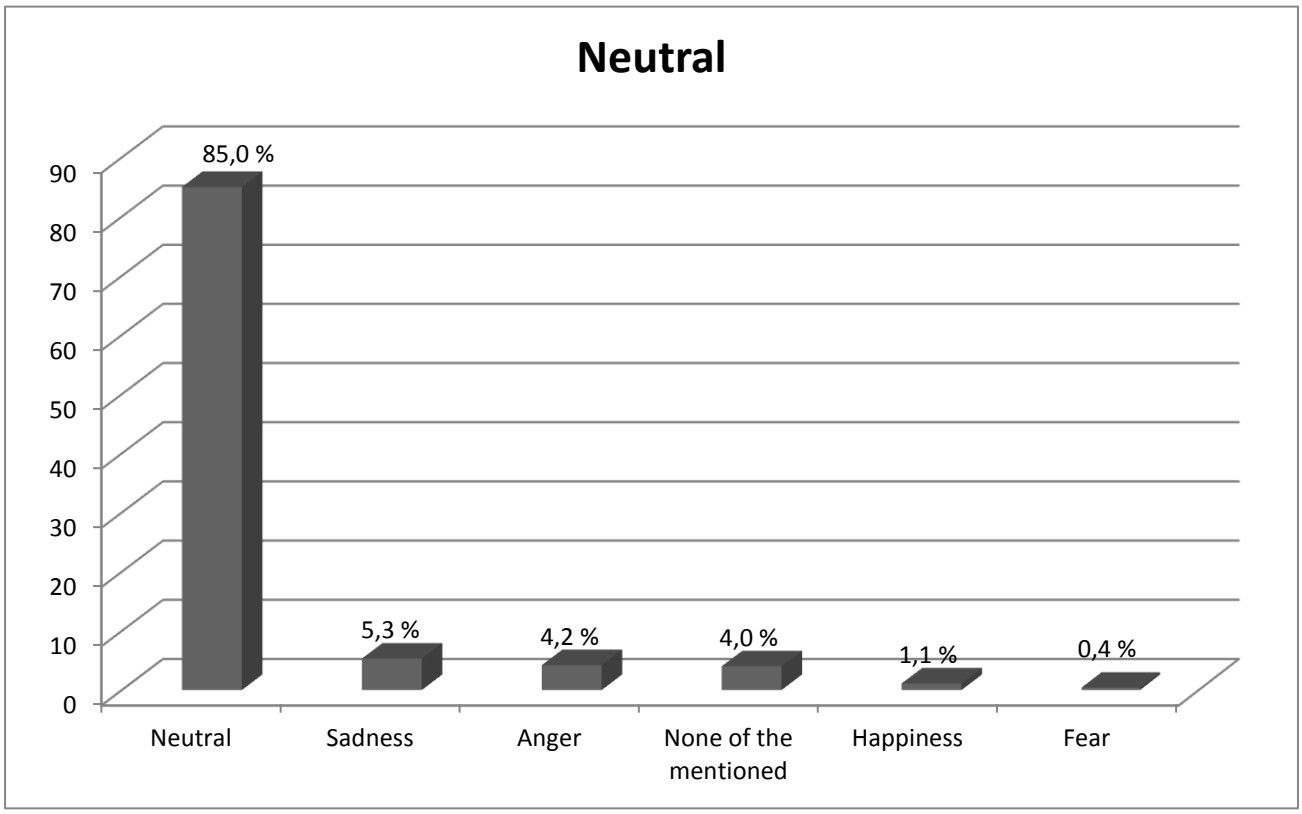


Figure 12: Chosen answer options for neutral faces in Emotion Recognition I, presented in percentages for the whole sample.

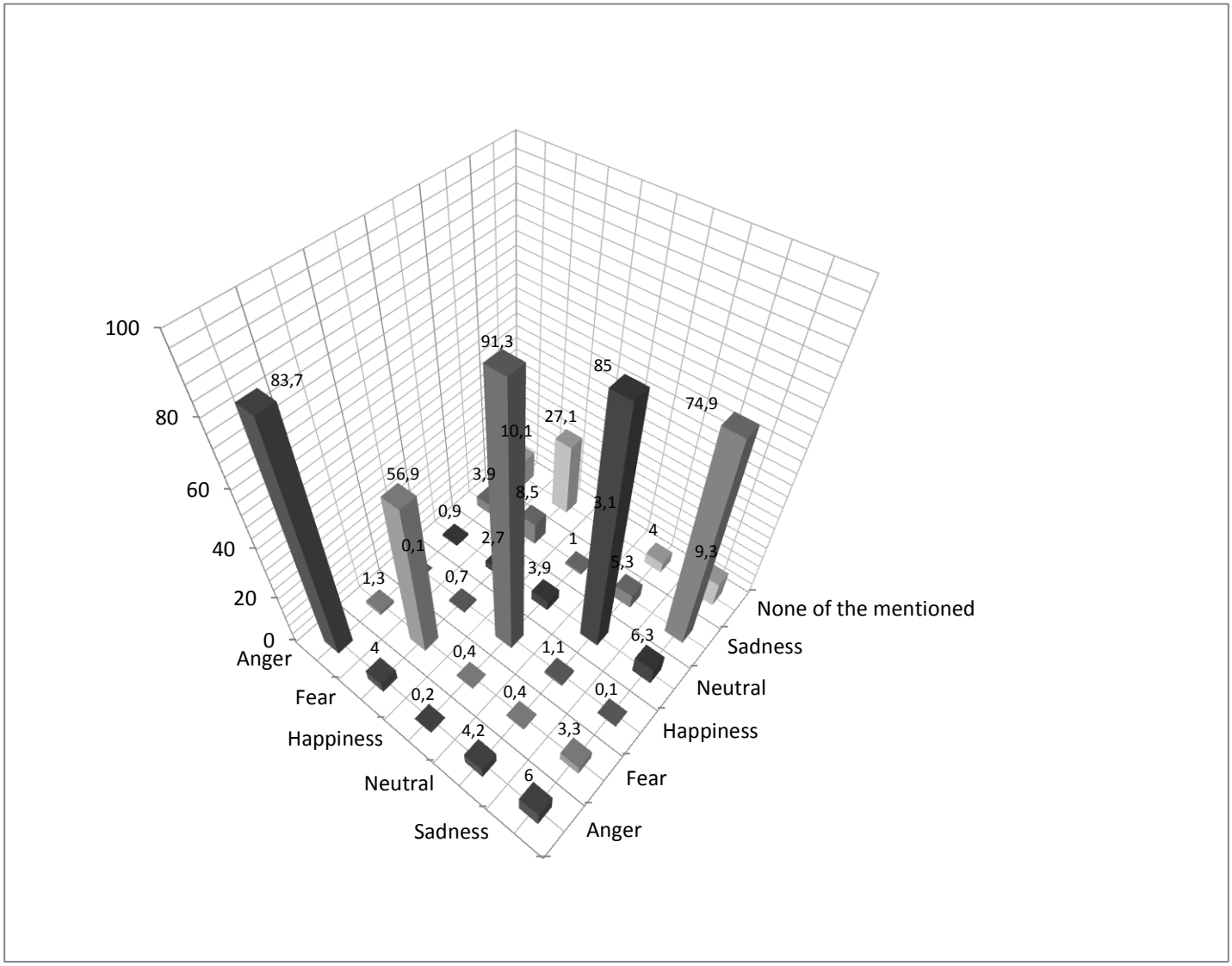


Figure 13: Chosen answer options for all emotions as well as neutral faces in Emotion Recognition I, presented in percentages for the whole sample.

Figure 14: Chosen answer options for emotions happiness and fear in Emotion Recognition II, presented in percentages for the whole sample.

H1(11): **There is a significant difference in overall facial emotion recognition accuracy rates regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two factorial ANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and overall recognition accuracy as dependent variable was applied to analyze differences in general emotion recognition between age groups and gender.

The ANOVA revealed a significant main effect for age ($F(2,84) = 7.909$, $p = 0.001$) but no significant main effect for gender ($F(1,84) = .073$, $p = .788$). Furthermore there was no significant interaction effect between age groups and gender ($F(2,84) = 2.629$, $p = .078$).

Post hoc analyses (Scheffé) revealed the young aged adults exhibited significantly higher results in overall recognition accuracy compared to middle aged ($p = .001$) and elderly adults which did not differ in recognition accuracy ($p = .217$), for an overview see table 7.

H1(12): **There is a significant difference in recognition accuracy concerning emotions anger, fear, happy and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and recognition accuracy rates for emotions anger, fear, happiness and sadness as well as for neutral faces as dependent variables was conducted.

The MANOVA showed significant main effects for age regarding the recognition of facially expressed fear ($F(2,84) = 5.135$, $p = .008$), sadness ($F(2,84) = 19.596$, $p \leq .0001$) and neutral faces ($F(2,84) = 3.132$, $p = .049$) and there were no significant main effects for angry ($F(2,84) = .891$, $p = .414$) and happiness ($F(2,84) = .662$, $p = .519$) recognition. Moreover there were no significant main effects for gender for the recognition of facially expressed anger ($F(1,84) = 3.291$, $p = .073$), fear ($F(1,84) = .477$, $p = .492$), happiness ($F(1) = .326$, $p = .569$), sadness ($F(1,84) = 1.633$, $p = .205$) and neutral faces ($F(1,84) = .02$, $p = .887$). There was a significant interaction effect between age and gender for recognition of happiness ($F(2,84) = 4.76$, $p = .011$). There were no significant interaction effects for the other emotions, namely for anger ($F(2,84) = .105$, $p = .9$), fear ($F(2,84) = .867$, $p = .424$) and sadness ($F(2,84) = .637$, $p = .532$) as well as for neutral faces ($F(2,84) = 1.354$, $p = .264$).

Post hoc analyses (Scheffé) showed that fear recognition was more accurate in the group of young adults than in the middle aged group ($p = .011$) but not compared to elderly adults ($p = .08$) whilst middle aged adults weren't more accurate than elderly adults ($p = 0.725$). With respect to sadness recognition young adults showed significantly higher accuracy rates compared to middle aged ($p = .001$) as well to elderly adults ($p \leq .0001$). Middle aged and elderly adults did differ significantly in the recognition of sadness ($p = .62$). Despite a significant main effect for gender concerning the recognition of neutral faces post hoc analyses (Scheffé) showed that middle aged adults only exhibited a tendency to recognize neutral faces more accurate than young adults ($p = .081$). Whilst middle aged adults did not

significantly differ from elderly aged adults (p = .131) regarding neutral face recognition also young adults and elderly adults show no significant differences (p = .973).

Whereas males showed lower accuracy rates (M = 13.4, SD = 1.12) than females (M = 14.27, SD = .79) in the young aged adults group regarding happiness recognition, males showed higher accuracy rates in the middle aged (M = 13.9, SD = .79) and elderly adult group (M = 13.8, SD = .67) than females (middle aged group: M = 13.1, SD = 1.68, elderly adults: M = 13.3, SD = 1.23), see figure 15 and table 7.

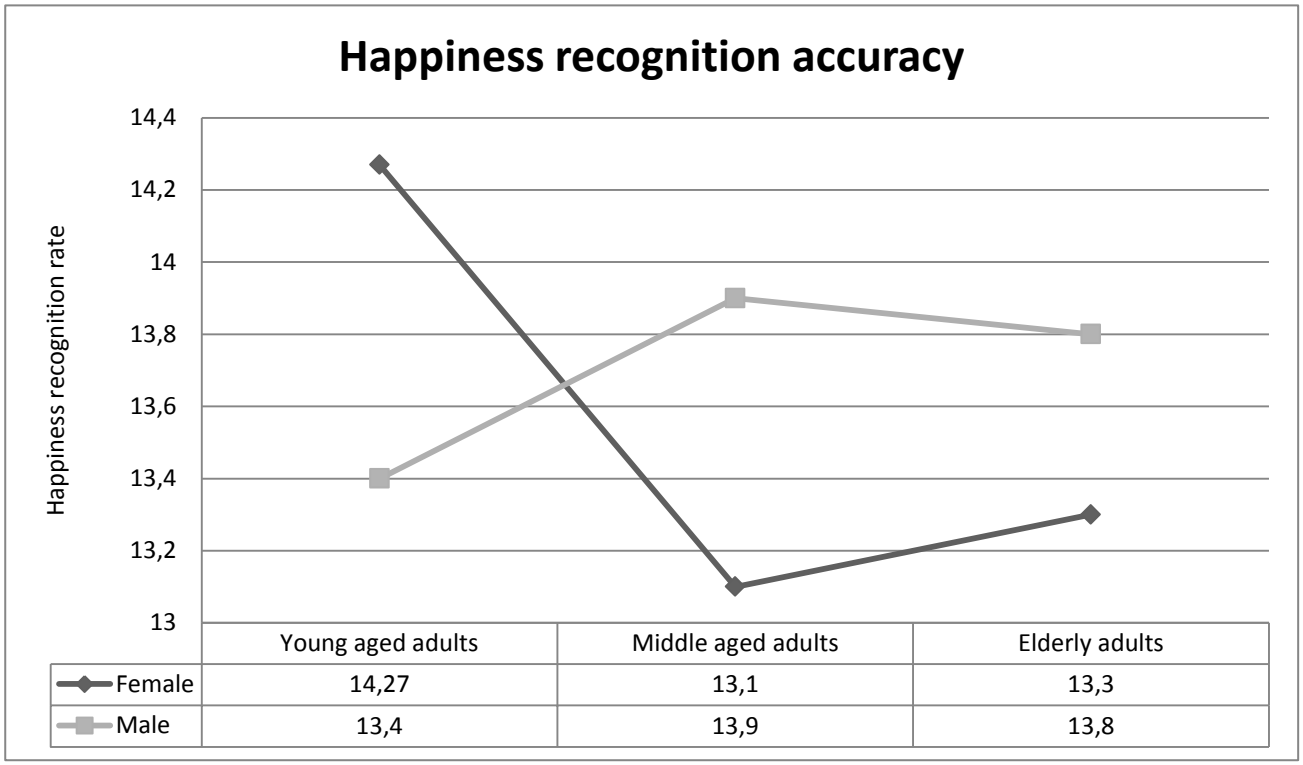


Figure 15: Gender and age interaction effects for happiness recognition accuracy scores in Emotion recognition I.

H1(13): **There is a significant difference in overall intensity rating of facially expressed emotions (no neutral faces included) regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two way factorial ANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and overall intensity rating as dependent variable was applied in order to analyze for differences in general intensity ratings.

There were no significant main effects for gender ($F(1,84) = .517$, $p = .474$) or age ($F(2,84) = .324$, $p = .725$), nor a significant interaction effect for gender and age ($F(2,84) = .82$, $p = .444$), for an overview see table 7.

H1(14): **There are significant differences in intensity ratings of facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and intensity ratings for facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces as dependent variables was calculated.

There were no significant main effects of age on intensity ratings of facially expressed emotions anger ($F(2,84) = .366$, $p = .694$), fear ($F(2,84) = .274$, $p = .761$), happiness ($F(2,84) = .803$, $p = .451$), sadness ($F(2,84) = 1.171$, $p = .315$) or neutral faces ($F(2,84) = .463$, $p = .631$). Furthermore there were no significant main effects for gender regarding intensity ratings of emotions anger ($F(1,84) = 3.002$, $p = .087$), fear ($F(1,84) = .394$, $p = .532$), happiness ($F(1,84) = .828$, $p = .366$), sadness ($F(1,84) = .249$, $p = .619$) as well as neutral faces ($F(1,84) = 1.925$, $p = .169$). Similarly there was no significant interaction effect for anger ($F(2,84) = .706$, $p = .496$), fear ($F(2,84) = .007$, $p = .993$), happiness ($F(2,84) = 2.674$, $p = .075$) or sadness ($F(2,84) = .649$, $p = .525$) intensity ratings as well as for neutral faces ($F(2,84) = 1.873$, $p = .16$).

H1(15): **There is a significant difference in overall arousal rating of facially expressed emotions (no neutral faces included) regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two way factorial ANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and overall arousal rating as dependent variable was applied in order to analyze for differences in general arousal ratings.

There were no main effects for gender ($F(1,84) \leq .0001$, $p = .991$) and age ($F(2,84) = .197$, $p = .822$), as well as no interaction effect for age and gender ($F(2,84) = 1.103$, $p = .337$), for an overview see table 7.

H1(16): **There are significant differences in arousal ratings of facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and arousal ratings for facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces as dependent variables was conducted.

There were no significant main effects of age on the arousal rating of facially expressed emotions anger ($F(2,84) = .445$, $p = 642$), fear ($F(2,84) = .104$, $p = .902$), happiness ($F(2,84) = .66$, $p = .52$) and sadness ($F(2,84) = .131$, $p = .877$) as well as for neutral faces ($F(2,84) = .946$, $p = .392$). Moreover there were no main effects for gender regarding arousal ratings for emotions anger ($F(1,84) = .596$, $p = .442$), fear ($F(1,84) = .157$, $p = .693$), happiness ($F(1,84) = 947$, $p = .333$), sadness ($F(1,84) = .059$, $p = .809$) or neutral faces ($F(1,84) = 2.506$, $p = .117$). Similarly there were no interaction effects for facially expressed anger ($F(2,84) = 1.383$, $p = .257$), fear ($F(2,84) = 1.385$, $p = .256$), happiness ($F(2,84) = .787$, $p = .459$), sadness ($F(2,84) = 1.141$, $p = .324$) or neutral faces ($F(2,84) = .503$, $p = .606$) regarding arousal ratings.

H1(17): **There is a significant difference in overall dominance rating of facially expressed emotions (no neutral faces included) regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two way factorial ANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and overall dominance rating as dependent variable was conducted so as to analyze for gender and age group differences in general arousal ratings.

There were no significant main effects for gender (F(1,84) = 1.655, p = .202) and age (F(2,84) = .389, p = .679) as well as no significant interaction effect of these two factors (F(2,84) = .843, p = 434) regarding overall dominance rating, for an overview see table 7.

H1(18): **There is a significant difference in dominance rating of facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and dominance ratings for facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces as dependent variables was calculated.

There were no significant main effects of age on dominance ratings of facially expressed anger (F(2,84) = 1.796, p = .172), fear (F(2,84) = .475, p = .624), happiness (F(2,84) = .416, p = .661),sadness (F(2,84) = .315, p = .731) or neutral faces (F(2,84) = .274, p = .761). Also there were no significant main effects for gender regarding dominance ratings of emotions anger (F(1,84) = .618, p = .434), fear (F(1,84) = 1.385, p = .243), happiness (F(1,84) = 1.053, p = .308), sadness (F(1,84) = .56, p = .573) as well as), neutral faces (F(1,84) = 3.006, p = .087). There was no significant interaction effect for emotions anger (F(2,84) = .708, p = .495), fear (F(2,84) = .891, p = .414), happiness (F(2,84) = .832, p = .439), sadness (F(2,84) = .56, p = .573) and for neutral faces (F(2,84) = 2.056, p = .134).

H1(19): **There is a no significant difference in recognition of facially expressed fear or happiness in the discrimination task (emotion recognition II) regarding age groups (young, middle and elderly adults) as well as gender (male, female).**

A two factorial MANOVA with age group (young, middle, elderly adults) and gender (male, female) as independent variables and accuracy rates for facially expressed fear and happiness as dependent variables was conducted.

There were no significant main effects of gender on recognition of facially expressed fear ($F(1) = .002$, $p = .969$) and happiness ($F(1,84) = .504$, $p = .48$) and no significant main effects of age on fear ($F(2,84) = 2.274$, $p = .109$) and happiness ($F(2,84) = 1.047$, $p = .355$) recognition. Moreover there was no significant interaction effect regarding gender and age group on fear ($F(2,84) = 1.206$, $p = .305$) or happiness ($F(2,84) = .48$, $p = .62$) recognition, for an overview see table 7.

H1(20): **There are significant correlations between the dimensions intensity, arousal and dominance concerning the facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces.**

Pearson correlations were conducted for each of the facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces separately in order to analyze the relationships between the rating dimensions.

For the emotion anger there was a significant coefficient of determinations between intensity and arousal ratings ($R^2 = .11$, $r = .326$, $p = .002$) whereas there were no significant coefficients of determinations for intensity and dominance ratings ($R^2 = .005$, $r = .069$, $p = .521$) as well as for arousal and dominance ratings ($R^2 = .0009$, $r = -.031$, $p = .769$).

For emotion fear there was a significant coefficient of determinations between intensity and arousal ratings ($R^2 = .21$, $r = .456$, $p = \leq .0001$ whereas there were no significant coefficients of determinations for intensity and dominance ratings ($R^2 = .006$, $r = -.078$, $p = .463$) as well as for arousal and dominance ratings ($R^2 = .005$, $r = -.069$, $p = .519$).

For emotion happiness there were significant coefficients of determinations between intensity and arousal ratings ($R^2 = .09$, $r = .293$, $p = .005$), between intensity and dominance ratings ($R^2 = .1$, $r = .316$, $p = .002$) as well as between arousal and dominance ratings ($R^2 = .02$, $r = .133$, $p = .002$).

For emotion sadness there was a significant coefficient of determinations for intensity and arousal ratings ($R^2 = .20$, $r = .452$, $p \leq .0001$). Furthermore there were no significant coefficients of determinations for the ratings of intensity and dominance ($R^2 = .000004$, $r = .002$, $p = .985$) as well as for arousal and dominance ($R^2 = .003$, $r = -.051$, $p = .636$).

For the neutral faces there was a significant coefficient of determinations for intensity and arousal ratings (R^2 = .05, r = .223, p = .035). There were no significant coefficients of determinations for the ratings of intensity and dominance (R^2 = .03, r = .168, p = .113) as well as for arousal and dominance (R^2 = .04, r = -.065, p = .544).

H1(21): **There is a significant correlation between the ratings of attractiveness and intensity of facially expressed emotions anger, fear, happiness and sadness as well as neutral faces.**

Pearson correlations were conducted for the attractiveness ratings and the intensity ratings of the facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces.

There was a significant coefficient of determinations between attractiveness ratings and intensity ratings of anger (R^2 = .04, r = .210, p = .047). There were no significant coefficients of determinations between the ratings of attractiveness and intensity ratings of fear (R^2 = .04, r = .199, p = .06), happiness (R^2 = .004, r = .064, p = .551) and sadness (R^2 = .02, r = .158, p = .136) as well as for neutral faces (R^2 = .0001, r = .01, p = .927).

H1(22): **There is a significant correlation between the ratings of attractiveness and arousal of facially expressed emotions anger, fear, happiness and sadness as well as neutral faces.**

Pearson correlations were conducted for the attractiveness ratings and the arousal ratings of the facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces.

There were significant coefficients of determinations between attractiveness ratings and arousal ratings of anger (R^2 = .04, r = .223, p = .035), fear (R^2 = .07, r = .262, p = .013) and happiness (R^2 = .07, r = .259, p = .014). There were no significant coefficients of determinations between the ratings of attractiveness and arousal ratings of sadness (R^2 = .04, r = .2, p = .58) and neutral faces (R^2 = .03, r = .178, p = .093).

H1(23): **There is a significant correlation between the ratings of attractiveness and dominance of facially expressed emotions anger, fear, happiness and sadness as well as neutral faces.**

Pearson correlations were conducted for the attractiveness ratings and the dominance ratings of the facially expressed emotions anger, fear, happiness and sadness as well as for neutral faces.

There were no significant coefficients of determinations between the ratings of attractiveness and dominance ratings of anger (R^2 = .02, r = .157, p = .141), fear (R^2 = .01, r = .106, p = .318), happiness (R^2 = .01, r = .116, p = .276) and sadness (R^2 = .02, r = .15, p = .159) as well as for neutral (R^2 = .02, r = .15, p = .159).

Table 7: Means (M) and Standard Deviations (SD) for the variables *emotion recognition I* (Overall Emotion Recognition Accuracy, Anger Recognition Accuracy, Fear Recognition Accuracy, Happiness Recognition Accuracy, Sadness Recognition Accuracy and Neutral Recognition Accuracy), *intensity rating* (Overall Intensity Ratings), *arousal rating* (Overall Arousal Ratings), *dominance rating* (Overall Dominance Ratings) and *emotion recognition II* – fear and happiness discrimination (Emotion Recognition II – Fear Accuracy and Recognition II – Happy Accuracy). Presented for the young adult group (Young), middle aged adult group (Middle) and elderly adults (Elderly) as well as for males and females and the gender-age-subgroups.

Means and Standard Deviations (N = 90)

Gender	Age Group		Overall Emotion Recognition Accuracy	Anger Recognition Accuracy	Fear Recognition Accuracy	Happiness Recognition Accuracy	Sadness Recognition Accuracy	Neutral Recognition Accuracy	Overall Intensity Ratings	Overall Arousal Ratings	Overall Dominance Ratings	Emotion Recognition II - Fear Accuracy	Emotion Recognition II - Happiness Accuracy
Male	Young	M	59.47	11.87	9.13	13.4	13.27	11.8	354.73	230.27	302.4	28.33	29.2
		SD	5.24	1.92	2.75	1.12	1.16	2.04	63.67	82.9	55.57	1.72	1.21
	Middle	M	59.2	12.53	7.73	13.93	11.53	13.47	337.8	233.33	284.73	29.07	29.47
		SD	5.58	1.55	3.19	.8	2.33	1.6	66.27	92.53	86.05	.7	.833
	Elderly	M	56.4	12	7.93	13.8	9.8	12.87	361.53	204.93	315.2	26.67	29.27
		SD	5.23	2.17	2.28	.68	2.73	2.23	73.77	99.16	93.91	3.87	1.22
	Overall	M	58.36	12.13	8.27	13.71	11.53	12.71	351.36	222.84	300.78	28.02	29.31
		SD	5.41	1.88	2.77	.89	2.56	2.05	67.24	90.61	79.18	2.62	1.08
Female	Young	M	63.8	12.87	10.66	14.27	13.4	12.6	346.53	230.13	305	28.4	28.73
		SD	5.52	2.61	3.56	.8	1.5	2.85	58.81	112.46	66.25	1.45	1.58
	Middle	M	57.4	13.27	7.2	13.13	10.27	13.53	371.07	195.47	331.8	27.93	29.27
		SD	6.18	1.28	2.91	1.69	2.9	1.36	64.84	66.35	73.9	2.94	1.22
	Elderly	M	54.87	12.53	8.27	13.3	8.93	11.8	365.73	242.2	324	27.67	29.4
		SD	7.14	2.03	3.43	1.23	3.45	2.78	57.52	123.445	43.9	3.68	.91
	Overall	M	58.69	12.89	8.71	13.58	10.87	12.64	361.11	222.6	320.27	28	29.13
		SD	7.25	2.02	3.55	1.36	3.28	2.48	60.04	103.32	62.26	2.8	1.27
Overall	Young	M	61.63	12.37	9.9	13.83	13.27	12.2	350.633	230.2	303.7	28.37	28.97
		SD	5.73	2.31	3.22	1.05	1.16	2.47	60.37	97.11	59.65	1.56	1.4
	Middle	M	58.3	12.9	7.47	13.13	10.9	13.5	354.43	214.4	308.27	28.5	29.37
		SD	5.86	1.45	3.01	1.68	2.66	1.47	66.6	81.42	82.37	2.17	1.03
	Elderly	M	55.63	12.27	8.1	13.57	9.37	12.33	363.63	223.57	319.6	27.17	29.33
		SD	6.2	2.08	2.87	1	3.09	2.54	65.03	111.64	72.16	3.74	1.06
	Overall	M	58.52	12.51	8.49	13.64	11.2	12.68	356.23	222.72	310.52	28.01	29.22
		SD	6.36	1.98	3.18	1.14	2.95	2.26	63.57	96.62	71.5	2.7	1.18

7. Interpretation

7.1 Cultural-, age-, and gender effects on the ratings of words' emotional properties

7.1.1 Rating words on their emotional properties and how cultural aspects influence these ratings

In the course of successfully conducting the Viennese Emotional Go/no-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) participants have to identify words as either being neutral or negative in nature in order to be able to adequately react to them or to inhibit a reaction. In order to present such words with valence ratings of either being negative or being neutral, words were chosen from the Berlin Affective Word List Reloaded - BAWL-R (Voe, et al., 2009) according to fixed selection criteria (previously described in section 3.3.5). The 100 negative words chosen for the Viennese Emotional Go/no-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) were significantly rated more negative than the chosen 100 neutral words, analyses relying on the normative ratings provided by the BAWL-R (Voe et al., 2009). Yet these results have to be seen critically due to missing sample specific information about nationality, native language or other cultural aspects (Voe, et al., 2009). Gole and colleagues (2012) used words taken from the BAWL-R for an emotional go/no-go task letting the participants additionally rate the words on valence and arousal showing that the previously classified negative words were rated significantly more negative than the neutral words. Beside these findings also Gole and colleagues (2012) do not provide any information about nationality, native language or other cultural aspects characterizing the sample. Despite a number of studies which show cross-cultural similarities regarding the rating of words' emotional connotations also studies can be found which show slight differences in word ratings indicating possible cultural specificities (for a discussion on this topic see section 3.3.2). These unclear results regarding universalities or specificities in word ratings imply the validation of the Viennese Emotional Go/no-go Task items (Willinger, Schmoeger, Deckert, & Auff, 2012) so as to obtain representative normative data for the chosen words.

In order to validate the words chosen for the Viennese Emotional Go/no-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) Computer Task I was conducted in a sample comprising of Austrians as well as German speaking people living in Austria, asking participants to rate a number of words on the dimensions valence, arousal, dominance, imagery and familiarity. First of all results showed that the previously classified neutral words were rated significantly more positive than the negative words. According to this result the words chosen for the Viennese Emotional Go/no-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) therefore can be divided into two classes of words: namely negative valenced words and such ones who are rated significantly more positive. Due to the fact that previously classified positive words weren't provided in the rating tasks it isn't admissible to label the words rated significantly more positive than the negative ones as 'neutral' words, strictly speaking. For the better understanding and because of the facts that in the Viennese Emotional Go/no-go Task (Willinger, Schmoeger, Deckert, & Auff, 2012) participants are explicitly instructed to distinguish between negative and neutral words as well as that these words (rated significantly more positive than negative words) were rated on average around the midpoint of the valence scale they will nevertheless be called neutral words. As negative words are rated significantly less positive than neutral words in this diploma thesis a result hopefully expected and fully in line with studies relying on the same words (Gole et al., 2012; Voe et al., 2009) could be shown. Regarding rating the words on arousal it could be shown that Austrians and German speaking people living in Austria rate negative words significantly more arousing than neutral words. This result is also equal to other studies showing negative words being generally rated as more arousing (Eilola & Havelka, 2010; Ferré et al., 2012; Gilet et al., 2012; Gole et al., 2012; Kanske & Kotz, 2010; Kanske & Kotz, 2011; Moors et al., 2013; Redondo et al., 2007; Soares et al., 2012; Soederholm et al., 2013; Voe, et al., 2009).

Beside these findings further analyses were conducted in order to look for culture-specific or cross-cultural similar patterns regarding relationships of the rating dimensions valence, arousal and dominance. For neutral words the correlation between ratings of valence and arousal was relatively low, a result that can also be found in other studies (e.g. Soares et al., 2012). For negative words on the other hand a medium-sized negative relationship ($r = -.538$) between valence and arousal ratings could be shown. This result is fully in line with studies commonly showing that lower valence ratings for negative words are accompanied with higher arousal ratings (Eilola & Havelka, 2010; Ferré et al., 2012; Gilet et al., 2012; Gole et al., 2012; Kanske & Kotz, 2010; Kanske & Kotz, 2011; Moors et al., 2013; Redondo et al.,

2007; Soares et al., 2012; Soederholm et al., 2013; Voe, et al., 2009). Dominance, being a dimension seemingly explaining less variance than valence and arousal (Bradley & Lang, 2000; Russell, 2003), did only show a low correlation with valence for neutral as well as for negative words. This result is somehow opposing other studies showing that negative words tend be rated as less controllable (Soares et al., 2012) or even showing strong positive relationships between dominance and valence ratings (Warinner et al., 2013). These inconsistent results could arise from the heterogeneous use of adjectives to describe the dimension dominance in instructions (Bradley & Lang, 1999; Redondo et al., 2007; Soares et al., 2012). Substantiating this assumption participants in this diploma thesis most often asked for further explanations of the dimension dominance compared to all other dimensions.

Thus can be shown that the representative sample in this diploma thesis shows comparable ratings and rating patterns which sustain the assumption of cross-cultural similar patterns regarding word ratings on the one hand but at the same time results are shown that could be due to culture- and/or language specific aspects.

7.1.2 Rating words on their emotional connotations whilst focusing on possible age- or gender effects

In search for possible age- or gender differences concerning the perception and processing of language and its emotional connotations three age groups with equal gender ratios conducted Computer Task I. Analyses revealed a significant effect for gender of the participant indicating that females rate negative words significantly more negative in the valence rating than males do. Whilst such gender differences in explicit valence rating haven't been found for adults yet this result additionally opposes the findings of Soares and colleagues (2012) which did not find gender differences in valence rating for adults at all. For the valence rating of neutral words no such gender differences could be found, a result that could be assumed due to the assumption that neutral words would be most likely rated within a narrow range around the midpoint of the valence scale. For the valence rating of neutral as well as negative words no differences could be found regarding age of participants. This result somehow contrasts to the studies of Gilet and colleagues (2012) as well as Gruhn & Smith (2008) showing age-related differences in words' valence rating. Beside the results concerning valence ratings further analyses were conducted in order to search for possible gender or age differences regarding the arousal rating of words. For the arousal ratings of negative as well

as neutral words no significant gender differences could be found albeit a tendency for females to rate negative words more arousing than males could be found. Although this gender difference did not reach significance it is nevertheless in line with other studies showing that females tend to rate negative words with higher arousal ratings than males do (Soares et al., 2012). For arousal ratings no age differences could be found.

7.2 **Cultural-, age-, and gender effects on the ratings of emotional and neutral faces**

7.2.1 The recognition of facially expressed emotions and neutral faces and its cultural aspects

In the course of this diploma thesis Computer Task II was conducted in order to validate the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012) as well as Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013) items, namely pictures of 15 different human photo models each one displaying angry, fearful, happy and sad expressions. These were randomly drawn from the Karolinska Directed Emotional Faces - KDEF (Lundqvist et al. 1998) and previously validated by Calvo and Lundqvist (2008) as well as Goeleven and colleagues (2008). Due to the unclear acquired body of knowledge regarding cross-cultural universality and specificity of human facial emotion perception and processing (extensively discussed in section 2.2.1) these pictures of facially expressed emotions had to be validated within a sample of Austrians as well as German speaking in Austria living people. This was done in order to obtain meaningful results considering the future usage of the Viennese Emotional Stroop and Viennese Affective Flexibility Tasks (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) and possible future modifications.

First of all it was important to show that these facially expressed emotions would be recognized as the ones meant to be displayed due to the fact that the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) performance depends on the capability of successfully recognizing the facially expressed emotions as either happy, fearful, angry or sad. To obtain results that would be representative for Austrians and people living in Austria and within its culture the task Emotion recognition I asked participants to recognize the displayed emotion

and choosing one of the given answer options. Besides the previously mentioned facially expressed emotions happiness, fear, anger and sadness also neutral faces had to be recognized in this emotion recognition task in order to obtain results concerning possible tendencies of labeling neutral expressions with other emotions. In general all expressions were recognized above the chance level of 16% with happiness being recognized most accurate (89%) followed by anger (84%), neutral (73%), sadness (60%) and fear being recognized least accurately (52%). This goes in line with other studies showing happy faces being recognized most accurately (Beaupré & Hess, 2005; Biehl et al., 1997; Calvo & Lundqvist, 2008; Ebner et al., 2010; Goeleven et al., 2008; Langner et al., 2010; Palermo & Coltheart, 2004; Tottenham et al., 2009) and fearful faces showing the lowest recognition accuracy (Beaupré & Hess, 2005; Biehl et al., 1997; Calvo & Lundqvist, 2008; Goeleven et al., 2008; Palermo & Coltheart, 2004; Tottenham et al., 2009) as well as in line with the studies which validated the KDEF faces (Calvo & Lundqvist, 2008; Goeleven et al., 2008).

Whereas for some expressions the percentage of wrongly chosen answers was very low e.g. happy faces which were sometimes confused with answers options 'neutral' (3.9%) or 'none of the mentioned' (3.1%) other showed higher confusion rates like e.g. anger being labeled with 'none of the mentioned' (10.1%) or 'sadness' (3.9%). For sad faces showing only low recognition accuracy at all (60%) it could be shown that the percentage of wrongly chosen answer options were "fairly" distributed showing confusion with 'none of the mentioned' (9.3%), 'neutral' (6.3%), 'anger' (6%) and 'fear' (3.3%). For fearful faces being recognized least accurately (only 52%) the answer distribution showed an obvious bias towards the answer option 'none of the mentioned' chosen in 27.1% of cases. Beside this bias fearful was most often confused with 'sadness' (8.5%) and 'anger' (4%). For neutral faces being accurately recognized in 73% of all cases it could be shown that neutral is most often confused with 'sadness' (5.3%), 'anger' (4.2%) and 'none of the mentioned'.

Due to the fact that surprised faces are not shown and therefore that specific emotion plays no (intended) role in the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) no faces displaying surprise were shown in the emotion recognition task and therefore the answer option 'surprise' wasn't provided. In the emotion recognition task fear recognition showed a severe answer bias towards the answer option 'none of the mentioned' (27.1% of the cases) and it is likely that participants did choose this answer as a substitute for surprise. Due to the fact that

participants weren't asked which answer they would have provided instead of choosing 'none of the mentioned' this can only be assumed but not be proofed. Nevertheless this assumption could be sustained by other studies showing that fear is predominantly confused with surprise (Calvo & Lundquist, 2008; Langner et al., 2010; Palermo & Coltheart, 2004; Tottenham et al., 2009). Beside this bias fear was most often confused with sadness (8.5%), which also goes in line with other studies showing similar confusion patterns (Calvo & Lundquist, 2008; Tottenham et al., 2009). Whilst the possible surprise colored choice of 'none of the mentioned' wouldn't pose such a problem for the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) due to the fact that surprised faces aren't presented or don't have to be sorted to respectively, the tendency to mistake fearful expressions for sad expressions would be more problematic. However an estimated probability of 8.1% confusing fear with sadness is not very likely to influence the performance in the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2013) in a significant way. That shouldn't attenuate the fact that surprise is least accurately recognized amongst all other displayed emotions in the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013). The specific process of recognizing fear compared to happiness will be discussed in section 7.2.3.

Despite sadness and fear showing unsatisfyingly low recognition accuracy rates the faces chosen for the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) were all recognized above the chance level and showed similar accuracy and confusion patterns like studies validating the same faces within samples of different nationalities and cultural aspects (Calvo & Lundqvist, 2008; Goeleven et al., 2008) as well as other studies validating other sets of emotional faces (Beaupré & Hess, 2005; Biehl et al., 1997; Ebner et al., 2010; Langner et al., 2010; Palermo & Coltheart, 2004; Tottenham et al., 2009). Thus can be summarized that the results from this diploma thesis are so far in line with those of other studies validating faces displaying emotional expressions indicating cross-culturally similar patterns of facial emotion perception and processing. Furthermore slightly higher anxiety characteristics do not seemingly influence emotion perception systematically. Despite fearful and sad faces showing low accuracy rates it could be shown that the chosen stimuli are most likely suitable to be used in the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) successfully applying them to Austrians and

German speaking people living in Austria. This can also be assumed due to the specific confusion patterns regarding fear and sadness recognition which potentially not seem to be influential within the specific requirements of the discussed emotion tasks.

7.2.2 Age and gender potentially influencing the recognition of facially expressed emotions and neutral faces

In order to make statements about the meaningful usage of the Viennese Emotional Stroop and Viennese Affective Flexibility Tasks (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) in different age and gender groups three age groups with equally distributed gender ratios conducted both computer tasks.

Regarding general emotion recognition accuracy (conducting Emotion recognition task I) it could be shown that young aged adults exhibited significantly better performances than middle aged and elderly adults. This is in line with other studies showing young aged adults being seemingly more accurate in overall emotion recognition than elderly adults (Ebner, et al., 2010; Horning, 2012; Leime, et al., 2013; Sasson et al., 2010; Suzuki & Akiyama, 2013; Weidner, 2014) and middle aged adults (Ebner, et al., 2010). Despite the results found in the course of this diploma thesis general conclusions about age differences in emotion recognition accuracy shouldn't be drawn carelessly due the aspect that the choice of emotions to be recognized differs slightly from study to study. Whereas an age effect on the overall recognition of emotions could be found the analyses did not reveal gender differences or age-gender interactions that would significantly influence general emotion recognition accuracy. These results are partly in line with other studies showing no gender differences in emotion recognition accuracy (Calvo & Lundquist, 2008; Hutchison & Gerstein, 2012, Melchers, et al., 2013) whilst it opposes other studies showing general advantages for females (Hall, 1978; Hall et al, 2000).

Focusing on the recognition of specific emotions rather than overall recognition accuracy interesting results could be found. Whilst for emotions anger, fear, happiness and sadness as well as for neutral faces being male or female did not have an influence on recognition accuracy, age did have a significant impact on the recognition of fear and sadness as well as a significant yet unclear impact on the processing of neutral faces. Recognizing facially expressed fear it could be shown that young aged adults exhibited more accurate ratings than

middle aged adults which themselves did not differ from elderly adults whilst maybe more interestingly there was no significant difference in recognition accuracy between young aged and elderly adults (albeit a trend could be seen benefitting young adults, $p = .08$). Somehow this result is somewhat contradictory to previous studies showing young adults outperforming elderly adults concerning fear recognition (Horning, 2012; Leime, et al., 2013; Suzuki & Akiyama, 2013). Recognizing facially expressed sadness young adults were significantly more accurate than middle aged as well elderly adults which did not substantially differ from each other. This is fully in line with other studies showing highest fear recognition accuracy rates for young adults (Ebner, et al., 2010; Horning, 2012; Leime, et al., 2013; Suzuki & Akiyama, 2013). Whilst analyses showed that there was a significant effect of age group on the recognition of neutral faces, post hoc test unfortunately did not show differences between age groups albeit interestingly a trend could be seen indicating middle aged adults being more accurate than young adults whilst not differing from elderly adults. This could be seen as a stronger tendency for young adults of labeling neutral faces with other emotions rather than assigning it no evolved emotion. Concerning the recognition accuracy of happy faces interesting age-gender effects could be found. Whilst some studies show that the recognition of facially expressed happiness seems to be uninfluenced by age differences (Ebner, et al., 2010; Ma, et al., 2013; West, et al., 2012) other studies show that age differences (Horning, 2012; Suzuki & Akiyama, 2013) as well as gender differences (de Carvalho Pinto et al., 2013) regarding happiness recognition are possible. Adding new results to the question of age- and gender influences on the recognition of happy faces it could be shown that whilst in the young adult group females are more accurate to recognize happy faces it looks like that this gender advantage seems to turn when getting older as for middle aged and elderly adults males seem to exhibit greater accuracy than females. Whilst males show an increase in happiness recognition accuracy from young adulthood to middle aged adulthood and then only slightly decrease on their way to elderly adulthood it seems that for females the aging pattern in happiness recognition seem to be opposite. Whilst young females exhibit best performances of all age-gender subgroups they seem to decrease in happiness recognition accuracy on their way to middle aged adulthood whilst improving again in elderly adulthood. Replicating these results and potentially finding more proof for such gender-age patterns regarding happiness recognition could be implications for future research.

Overall could be shown that age and gender seem to be potential factors influencing the recognition of facially expressed emotions whilst the fact that results could be found being

consistent as well as inconsistent with other studies concerning facial emotion recognition future research on this topic is implied.

7.2.3 Discriminating fearful and happy facial expressions

In order to validate the Viennese Emotional Stroop Task items (Willinger, Schmoeger, Deckert, & Auff, 2012) participants had to conduct an additional emotion recognition/discrimination task namely Emotion recognition II. Due to the fact that in the Viennese Emotional Stroop Task (Willinger, Schmoeger, Deckert, & Auff, 2012) solely happy and fearful faces are presented participants had to discriminate between faces either displaying happy or fearful expressions and pick the answer options ‘fear’ or ‘happiness’ in a forced choice task.

Despite the low recognition accuracy for fearful faces in Emotion recognition task I participants scored high albeit not perfect recognition scores in Emotion recognition task II. With recognition accuracy rates of 92% for fearful faces as well as 93% for happy faces both expressions were recognized far beyond the chance level of 50%. Despite high accuracy rates for happy faces in general as well as happy and fearful faces very seldom confused for one another (Beaupré & Hess, 2005; Biehl et al., 1997; Calvo & Lundqvist, 2008; Ebner et al., 2010; Goeleven et al., 2008; Langner et al., 2010; Palermo & Coltheart, 2004; Tracy et al, 2009; Tottenham et al., 2009) the results of happy faces being given the answer option ‘fear’ in 1.6% of all cases and especially fearful faces being labeled with ‘happiness’ in 6% of all cases have to be to critically noted. These unexpectedly high confusion rates for happy but especially for fearful faces are most likely due to decreasing attention capacities of participants in combination with conducting Emotion recognition task II too fast at the end of an approximately two hour taking study participation. These accuracy decreases could also have been potentiated by the switching answer positions of ‘happiness’ and ‘fear’. Further analyses showed that there were no age-, gender or interaction effects regarding the accuracy of recognizing happy or fearful faces when the task is to discriminate between those two expressions.

7.2.4 Age and gender influencing the ratings of emotional faces?

In order to analyze for possible influential effects concerning the question how emotional as well as neutral faces are rated female as well as male participants of different age groups had to rate the faces they saw according to the dimensions intensity, arousal and dominance. This was done in order to obtain results whether amongst Austrians and German speaking people living in Austria there would occur specific age or gender differences rating the faces chosen for the Viennese Emotional Stroop Viennese and Affective Flexibility Tasks (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) according to dimensions typically asked for when validating faces expressing different emotions. Comparing age and gender groups no significant differences in overall intensity ratings, intensity ratings of specific emotions, overall arousal ratings, arousal ratings of specific emotions, overall dominance ratings as well dominance ratings for specific emotions couldn't be found at all. Regarding intensity ratings Beaupré and Hess (2005) as well as Biehl and colleagues (1997) indicated that cross-cultural differences seem to exist whilst other studies hint that age- and gender differences cause only slight variations concerning intensity ratings (Hutchison & Gerstein, 2012; Langer et al., 2010). Similarly to these results it could be shown that intensity ratings did not differ significantly between age- and gender groups which has to be noted positively regarding the future usage of the Viennese Emotional Stroop Viennese and Affective Flexibility Task items (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) heeding studies which hint that females seem to be superior when it comes to the recognition of only subtle facial expressions. Such differences would most likely influence the performance on the Viennese Emotional Stroop and Viennese Affective Flexibility Tasks (Willinger, Schmoeger, Deckert, & Auff, 2012, 2013) but due to the results such cofounding effects aren't very likely. Due to previous studies (Gruhn & Scheibe, 2008) indicating that age related differences seem to exist when rating pictures of different valence on arousal some studies control for arousal ratings in order to obtain more meaningful results concerning age differences (Czwerwon, Luettke, & Werheid, 2011). Opposing such results and assumptions no differences in arousal ratings between age- as well as gender groups could be found in this validation study. In comparison with studies finding age differences in arousal ratings of emotional pictures presenting different contents like the IAPS pictures (Gruhn & Scheibe, 2008) emotional faces seem to present an own class of objects which not seem to be processed equally. For dominance being a dimension usually not analyzed for when rating emotional faces no age- or gender differences could be found in

rating patterns maybe because dominance seems to be a dimension explaining less variance than valence or arousal (Bradley & Lang, 2000; Russell, 2003).

7.2.5 Relationships between ratings dimensions rating emotional faces

In order to analyze for specific relationships between the rating dimensions intensity, arousal, dominance and attractiveness Pearson correlations for each emotion separately were conducted. Dominance, usually not analyzed for when rating faces displaying emotional expressions, did not show mentionable relationships with the other rating dimensions intensity, attractiveness or arousal at all. Intensity and arousal however showed small till middle sized relationships for the emotions anger, fear and sadness whilst interestingly not for happiness or neutral faces. These results are partly in line with the findings of Goeleven and colleagues (2008) showing a high overall correlation between intensity and arousal in the course of rating emotional faces. For the negative valenced emotions anger, fear and sadness it seems that judgments how intense an expression is displayed and degrees of personal inner arousal and excitement seem to be at least moderately explaining variance for each other. This indicates that the more emotions of anger, fear and sadness are judged to be evolved the more inner arousal they trigger or vice versa. Similar relationships could not be found for emotion happy as well as for neutral faces, indicating that arousal and intensity ratings seem to be more independent for these specific expressions. In the style of Langner and colleagues (2010) participants had furthermore to rate the faces on the dimension attractiveness. This was done so as to analyze for attractiveness of a face being a potential factor influencing the general rating of emotional faces. It could be shown that attractiveness didn’t show mentionable relationships with the dimensions intensity, arousal or dominance indicating that attractiveness of faces is likely to be not influential regarding these rating dimensions seemingly being a dimension processed (mostly) independently.

8. Conclusion

In the years 2012 and 2013 the 'Viennese Emotional Test Battery' (VETeBA) was designed consisting of three tasks namely the Viennese Emotional Stroop Task, the Viennese Affective Flexibility Task and the Viennese Emotional Go/no-go Task (Willinger, Schmoeger and Deckert, & Auff, 2012, 2013) each task measuring different aspects of affective information processing. The performance in the VETeBa depends amongst other requirements on the ability to recognize human facially expressed expressions as well as distinguish between negative as well as neutral words. For the construction of the VETeBa pictures of human faces displaying specific emotions as well words with different emotional properties were randomly drawn from standardized databases (KDEF - Lundqvist et al. 1998; BAWL-R – Voe et al., 2009). An unclear body of knowledge in research regarding the existence of cross-cultural universalities and/or specificities concerning the perception and processing of facially expressed emotions as well as emotional verbal material leads to the necessity to validate these chosen faces and words within a sample representative for the population in which meaningful results conducting the VETeBa should be obtained. In order to do so two computer tasks were conducted in a sample of Austrians and German speaking people living in Austria which was divided into three age groups (young, middle and elderly adults) with equal numbers of males and females. This was done so as to validate the emotional material according to cross-cultural aspects as well as collect normative data of different age and gender groups in order to analyze whether the usage of the stimuli chosen for the VETeBa could potentially disadvantage one of these groups. In Computer task I participants had to rate the words chosen for the Viennese Emotional Go/no-go task on valence, arousal, dominance, imagery and familiarity. Results showed that the previously chosen neutral words from the BAWL-R were rated significantly more positive than the chosen negative words, showing that these two classes of words were clearly distinguished in this representative sample. Due to this result it is very likely that Austrians as well as German speaking people living in Austria should be able to recognize negative and neutral words whilst conducting the Viennese Emotional Go/no-go task. Whilst for the ratings of the chosen words no age differences could be found it could be shown that females rate negative words significantly more negative than males do whereas for arousal ratings females showed a tendency to associate negative words with higher arousal ratings than males. For the ratings of valence and arousal only for negative words a nearly medium-sized negative correlation could be found which is in line with the majority of studies conducting word ratings. In Computer Task II participants had to

validate pictures of human faces displaying different expressions chosen for the Viennese Emotional Stroop Task as well as the Viennese Affective Flexibility Task. Participants had to recognize the facial expressions in two different emotion recognition tasks as well as rate them on the dimensions attractiveness, intensity, arousal and dominance. In the first recognition task implying a multiple choice format all expressed emotions as well as the neutral faces were recognized above the chance level with happiness being recognized most accurately (89%) and fear least accurately (52%). With fear being confused with sadness in 8.5% of all cases this poses a slight problem for the Viennese Affective Flexibility albeit fortunately the magnitude of this problem is most likely not exceedingly large. This diploma thesis furthermore showed that young aged adults exhibit best performances in overall emotion recognition accuracy as well as for the particular emotions fear and sadness. Despite no gender differences for the whole sample regarding recognition accuracy could be found a significant age-gender interaction effect could be shown for the recognition of happy faces. Whilst in the young adult group females outperform males and show greatest accuracy of all subgroups in happiness recognition this advantage was shown to decrease with aging, showing that in the middle aged and elderly adult group males show higher accuracy than females. In the second recognition task participants had to recognize fearful as well as happy faces and label them with the answers ‘fear’ or ‘happiness’ in a forced choice task. Analysis showed high accuracy rates for fearful (92%) and happy faces (93%) far above the chance level albeit the confusion rates especially for fear were unexpectedly high most likely due to attention deficits at the end of the whole rating process. This assumption is substantiated by viewing the confusion rates of these two emotions in recognition task I. Nevertheless the recognition rates are most likely high enough to enable a successful execution of the Viennese Emotional Stroop task in the previously mentioned population.

Summarizing the results of this validation study both the verbal material and well as the pictorial material chosen for the VETeBa tasks can be characterized by the supposed emotional properties even within a population of Austrians as well as German speaking people living in Austria as normative data was obtained by a representative sample. Thus these faces and words can be seen as appropriate stimuli for the VETeBa tasks.

9. Conclusion German

In den Jahren 2012 und 2013 wurde die ‚Viennese Emotional Test Battery' – VETeBa (Wiener Testbatterie zur Erfassung affektiver Informationsverarbeitung) von Willinger, Schmöger, Deckert und Auff entwickelt. Diese Testbatterie setzt sich zusammen aus dem ‚Viennese Emotional Go/no-go Task' (Emotionaler Go/No-go), dem ‚Viennese Emotional Stroop' (Emotionaler Stroop) und dem ‚Viennese Affective Flexibility Task' (Affektive Flexibilität). Diese Testbatterie erfasst verschiedene Aspekte affektiver Informationsverarbeitung, indem die drei Tests nach verschiedenen Paradigmen konstruiert wurden und verschiedenartige emotionale Stimuli eingesetzt werden. Voraussetzungen für die erfolgreiche Durchführung der VETeBa-Tests sind auf der einen Seite das Erkennen von Emotionen im menschlichen Gesicht (Emotionen Angst, Freude, Trauer und Ärger) und auf der anderen Seite das Unterscheiden von negativen und neutralen Worten. Die emotionalen Gesichter als auch die emotional gefärbten Worte wurden zwar standardisierten Datenbanken (Lundqvist, Flykt, & Öhman, 1998; Vö et al., 2009) entnommen, jedoch ist die Notwendigkeit, dieses Material anhand repräsentativer Stichproben zu validieren, gegeben. Diese Notwendigkeit resultiert aus der Forschung zur menschlichen Emotionsverarbeitung, welche durchwegs uneinheitliche Resultate hervorbringt. So scheint es, dass es sowohl bei der Erkennung und Verarbeitung von Emotionen im menschlichen Gesicht, als auch bei der Verarbeitung emotionaler Sprachinhalte Ergebnisse gibt, die von kulturell universalen und somit gleichartigen Verarbeitungsweisen berichten, als auch Ergebnisse, die dafür sprechen, dass diese Vorgänge von Kultur zu Kultur unterschiedlich sind. Somit war es notwendig, die ausgewählten emotionalen Materialien anhand einer Stichprobe zu validieren, welche für in Österreich lebende Menschen repräsentativ ist, um die Ergebnisse der VETeBa-Tests sinnvoll interpretieren zu können. Zu diesem Zweck wurde eine Stichprobe von 90 gesunden Personen, welche sich aus drei Altersgruppen zusammensetzte (junge Erwachsene 20-30 Jahre, Personen mittleren Alters 40-50 Jahre und ältere Erwachsene 60-70 Jahre), gebeten, diese Stimuli zu bewerten. Zu diesem Zweck wurden zwei verschiedene computerisierte Verfahren mit dem Programm E-Prime 2.0 © programmiert, die den Probanden nacheinander vorgegeben wurden. Im ersten Computerverfahren mussten Personen die Worte nach ihrer Valenz bewerten, d.h. wie positiv oder negativ ein gegebenes Wort empfunden wird (weitere Bewertungskategorien waren Erregung, Dominanz, Mentale Vorstellung und Vertrautheit). Im zweiten Computerverfahren hatten die Probanden die Aufgabe im menschlichen Gesichtern dargestellte Emotionen (Angst, Freude, Trauer und Ärger) zu erkennen und diese

zusätzlich nach weiteren Kategorien zu bewerten (Attraktivität, Erregung, Dominanz und Intensität). Die Resultate zeigten, dass die im Vorfeld als negativ klassifizierten Worte signifikant negativer bewertet wurden, als die im Vorfeld als neutral klassifizierten Worte. Das bedeutet, dass die ursprüngliche, anhand der Datenbankwerte vorgenommene, Einteilung in negative und positive Worte auch für in Österreich lebende Personen Gültigkeit aufweist. Die Bewertung der Gesichter brachte hervor, dass alle Emotionen mit einer über dem Wahrscheinlichkeitslevel liegenden Genauigkeit bewertet wurden und auch die Verwechslungsraten in Hinblick auf die VETeBa-Tests günstige Ergebnisse lieferten. Somit konnte auch für die Gesichter gezeigt werden, dass die darin gezeigten Emotionen auch von in Österreich lebende Personen als diejenigen erkannt werden, die dargestellt werden sollen. Alters- und geschlechtsspezifische Aspekte in Bezug auf die Verarbeitung von emotionalen Gesichtern und Worten wurden untersucht und werden in dieser Diplomarbeit ausführlich berichtet.

Zusammenfassend lässt sich sagen, dass die ausgewählten Materialien anhand der im Vorfeld erwarteten emotionalen Eigenschaften bewertet wurden und somit ihr Einsatz im Zuge der VETeBa-Tests der Resultate dieser Validierungsstudie nach gerechtfertigt ist.

Abstract

In 2012/2013 the ‘Viennese Emotional Test Battery’ (VETeBa) was designed to assess different aspects of affective information processing presenting emotional faces and emotional words. Due to inconsistent results regarding universality of emotion perception these stimuli had to be validated within a sample representative for people living in Austria including 90 young, middle and elderly female and male adults. Conducting two computerized validation tasks it could be shown that the faces and words were adequately characterized by the supposed emotional properties and thus represent appropriate stimuli for the VETeBa. Age and gender aspects regarding facial and linguistic emotion perception are discussed.

Abstract (German)

In den Jahren 2012 und 2013 wurde die ‚Wiener Testbatterie zur Erfassung affektiver Informationsverarbeitung' entwickelt. Um verschiedene Aspekte dieser speziellen Art der Informationsverarbeitung zu erfassen, verwendet diese Testbatterie emotionale Gesichter und emotionale Wörter als Stimulusmaterial. Aufgrund inkonsistenter Forschungsergebnisse bezüglich kulturübergreifender Gemeinsamkeiten in der Emotionsverarbeitung ist es unumgänglich solche emotionalen Stimuli anhand einer, für in Österreich lebende Menschen, repräsentativen Stichprobe zu validieren. Die Stichprobe umfasste 90 Personen und setzte sich aus jungen Erwachsenen, Personen mittleren Alters und älteren Erwachsenen zusammen. Anhand der Ergebnisse zweier computerisierter Validierungsverfahren konnte gezeigt werden, dass sowohl die Gesichter als auch die Wörter anhand der im Vorfeld erwarteten emotionalen Eigenschaften charakterisiert werden konnten und somit als geeignetes Stimulusmaterial für die ‚Wiener Testbatterie zur Erfassung affektiver Informationsverarbeitung' betrachtet werden kann. Alters- und Geschlechtseffekte bezüglich gesichtsbezogener und sprachlicher Emotionsverarbeitung werden angeführt und diskutiert.

References

Abbassi, E. (2012). Affective word priming in the left and right visual fields in young and older individuals. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 72*(9-B), 5585.

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews, 1*(1), 21-62. doi: 10.1177/1534582302001001003

Andrade, N. C., Abreu, N. S., Duran, V. R., Veloso, T. J., & Moreira, N. A. (2013). Recognition of facial expressions of emotions: Standardization of pictures for Emotion Matching Tasks. *Psico, 44*(3), 382-390.

Anguas-Wong, A. M., & Matsumoto, D. (2007). Acknowledgement of emotional facial expression in Mexican college students. *Revista de Psicologia, 25*(2), 277-293.

Ashley, V., & Swick, D. (2009). Consequences of emotional stimuli: Age differences on pure and mixed blocks of the emotional Stroop. *Behavioral and Brain Functions, 5*, 14. doi: http://dx.doi.org/10.1186/1744-9081-5-14

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods Instruments & Computers, 34*(3), 424-434. doi: 10.3758/bf03195471

Barcelo, F. (2003). The Madrid card sorting test (MCST): a task switching paradigm to study executive attention with event-related potentials. *Brain Research Protocols, 11*(1), 27-37. doi: 10.1016/s1385-299x(03)00013-8

Bauer, L. M., & Altarriba, J. (2008). An investigation of sex differences in word ratings across concrete, abstract, and emotion words. *The Psychological Record, 58*(3), 465-474.

Beaupré, M. G., Cheung, N., & Hess, U. (2000). The Montreal Set of Facial Displays of Emotion [Slides].

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology, 36*(3), 355-370.

Bellace, M., Williams, J. M., Mohamed, F. B., & Faro, S. H. (2012). An fMRI study of the activation of the hippocampus by emotional memory. *International Journal of Neuroscience, 123*(2), 121-127.

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., et al. (1997). Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior, 21*(1), 3-21.

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale - An updated literature review. *Journal of Psychosomatic Research, 52*(2), 69-77. doi: 10.1016/s0022-3999(01)00296-3

Bourne, V. J., & Vladeanu, M. (2013). Examining the relationship between lateralisation for processing emotional faces, depression, and sex. *Laterality: Asymmetries of Body, Brain and Cognition, 18*(6), 748-766.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59.

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings (Tech.Rep. No. C-1)*. Gainesville: University of Florida, Center for Research in Psychophysiology.

Bradley, M. M., & Lang, P. J. (1999b). *International Affective Digitized Sounds (IADS): Stimuli, instruction manual and affective ratings*. Center for Research in Psychophysiology, Gainesville: University of Florida.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling and physiology. In R. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). New York: Oxford University Press.

Breitberg, A., Drevets, W. C., Wood, S. E., Mah, L., Schulkin, J., Sahakian, B. J., & Erickson, K. (2013). Hydrocortisone infusion exerts dose- and sex-dependent effects on attention to emotional stimuli. *Brain and Cognition, 81*(2), 247-255.

Briceno, E. M. (2014). The influence of gender and aging on the neural circuitry supporing facial emotion processing in adults with major depressive disorder. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74*(7-B(E)), No Pagination Specified.

Brickman, A. M., Zimmerman, M. E., Paul, R. H., Grieve, S. M., Tate, D. F., Cohen, R. A., . . . Gordon, E. (2006). Regional white matter and neuropsychological functioning across the adult lifespan. *Biological Psychiatry, 60*(5), 444-453. doi: 10.1016/j.biopsych.2006.01.011

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion norms for nouns: Berlin affective word list (DENN-BAWL). *Behavior Research Methods, 43*(2), 441-448. doi: 10.3758/s13428-011-0059-y

Burianova, H., Lee, Y., Grady, C. L., & Moscovitch, M. (2013). Age-related dedifferentiation and compensatory changes in the functional network underlying face processing. *Neurobiology of Aging, 34*(12), 2759-2767.

Busche, L. K. (2014). Explaining the cross-race effect: How inconsistency influences facial recognition. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74*(7-B(E)), No Pagination Specified.

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods, 40*(1), 109-115. doi: 10.3758/brm.40.1.109

Cao, R., Wang, S., Rao, C., & Fu, J. (2013). Task-irrelevant own-race faces capture attention: Eye-tracking evidence. *Scandinavian Journal of Psychology, 54*(2), 78-81.

Carre, J. M., Fisher, P. M., Manuck, S. B., & Hariri, A. R. (2012). Interaction between trait anxiety and trait anger predict amygdala reactivity to angry facial expressions in men but not women. *Social Cognitive and Affective Neuroscience, 7*(2), 213-221.

Castro, N., & James, L. E. (2014). Differences between young and older adults' spoken language production in descriptions of negative versus neutral pictures. *Aging, Neuropsychology, and Cognition, 21*(2), 222-238.

Chen, J. (2014). Face recognition as a predictor of social cognitive ability: Effects of emotion and race on face processing. *Asian Journal of Social Psychology, 17*(1), 61-69.

Chiu, P. H., Holmes, A. J., & Pizzagalli, D. A. (2008). Dissociable recruitment of rostral anterior cingulate and inferior frontal cortex in emotional response inhibition. *Neuroimage, 42*(2), 988-997.

Coney, J., & Fitzgerald, J. (2000). Gender differences in the recognition of laterally presented affective nouns. *Cognition and Emotion, 14*(3), 325-339.

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods Instruments & Computers, 36*(3), 384-387. doi: 10.3758/bf03195585

Czerwon, B., Luettke, S., & Werheid, K. (2011). Age Differences in Valence Judgments of Emotional Faces: The Influence of Personality Traits and Current Mood. *Experimental Aging Research: An International Journal Devoted to the Scientific Study of the Aging Process, 37* (5), 503-515. DOI: 10.1080/0361073X.2011.619468

Derntl, B., Habel, U., Robinson, S., Windischberger, C., Kryspin-Exner, I., Gur, R. C., & Moser, E. (2012). Culture but not gender modulates amygdala activation during explicit emotion recognition. *BMC Neuroscience, 13,* 54. doi: http://dx.doi.org/10.1186/1471-2202-13-54

De Carvalho Pinto, B. M., Dutra, N. B., Filgueiras, A., Juruena, M. F. P., & Stingel, A. M. (2013). Gender differences among undergraduates in the recognition of emotional facial expressions. *Avances en Psicologia Latinoamericana, 31*(1), 200-222.

Dewaele, J.-M., & Pavlenko, A. (2002). Emotion vocabulary in interlanguage. *Language Learning, 52*(2), 263-322.

Ding, X. P., Fu, G., & Lee, K. (2014). Neural correlates of own- and other-race face recognition in children: A functional near-infrared spectroscopy study. *Neuroimage, 85*(Part 1), 335-344.

Donges, U. S., Kersting, A., & Suslow, T. (2012). Women's Greater Ability to Perceive Happy Facial Emotion Automatically: Gender Differences in Affective Priming. *Plos One, 7*(7).

Dunajska, M., Szymanik, A., & Trempala, J. (2012). Attentional bias and emotion in older adults: Age-related differences in responses to an emotional Stroop task. *Polish Psychological Bulletin, 43*(2), 86-92.

Dundas, E. M., Plaut, D. C., & Behrmann, M. (2013). The joint development of hemispheric lateralization for words and faces. *Journal of Experimental Psychology: General, 142*(2), 348-358.

Ebner, N. C., He, Y., & Johnson, M. K. (2011). Age and emotion affect how we look at a face: Visual scan patterns differ for own-age versus other-age emotional faces. *Cognition and Emotion, 25*(6), 983-997.

Ebner, N. C., Johnson, M. K., & Fischer, H. (2012). Neural mechanisms of reading facial emotions in young and older adults. *Frontiers in Psychology, 3*, 223.

Ebner, N. C., Johnson, M. R., Rieckmann, A., Durbin, K. A., Johnson, M. K., & Fischer, H. (2013). Processing own-age vs. other-age faces: Neuro-behavioral correlates and effects of emotion. *Neuroimage, 78*, 363-371.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES - A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*(1), 351-362. doi: 10.3758/brm.42.1.351

Egner, T., Etkin, A., Gale, S., & Hirsch, J. (2008). Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. *Cerebral Cortex, 18*(6), 1475-1484. doi: 10.1093/cercor/bhm179

Eilola, T. M., & Havelka, J. (2010). Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods, 42*(1), 134-140. doi: 10.3758/brm.42.1.134

Ekman, P., (1972). Universals and cultural differences in facial expressions of emotion. . In J. Cole (Ed.), *Nebraska Symposium on Motivation (Vol. 19)* (pp. 207–282). Lincoln: University of Nebraska Press.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*(4), 384-392.

Ekman, P. (1992). Are the basic emotions? *Psychological Review, 99*(3), 550-553.

Ekman, P. (1999). Facial Expressions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. New York: John Wiley & Sons

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System*. Salt Lake City,UT: A Human Face.

Ekman, P., & Matsumoto, D. (2011). Reading faces: The universality of emotional expression *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 140-146). New York, NY: Worth Publishers; US.

Elfenbein, H. A. (2013). Nonverbal Dialects and Accents in Facial Expressions of Emotion. *Emotion Review, 5*(1), 90-96. doi: 10.1177/1754073912451332

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 203-235.

Elfenbein, H. A., Beaupre, M., Levesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion, 7*(1), 131-146. doi: 10.1037/1528-3542.7.1.131

Engelmann, J. B., & Pogosyan, M. (2013). Emotion perception across cultures: The role of cognitive mechanisms. *Frontiers in Psychology, 4*, 118.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron, 51*(6), 871-882. doi: 10.1016/j.neuron.2006.07.029

Etkin, A., Prater, K. E., Hoeft, F., Menon, V., & Schatzberg, A. F. (2010). Failure of Anterior Cingulate Activation and Connectivity With the Amygdala During Implicit Regulation of Emotional Processing in Generalized Anxiety Disorder. *American Journal of Psychiatry, 167*(5), 545-554. doi: 10.1176/appi.ajp.2009.09070931

Ferré, P., Guasch, M., Moldovan, C., & Sanchez-Casas, R. (2012). Affective norms for 380 Spanish words belonging to three different semantic categories. *Behavior Research Methods, 44*(2), 395-403. doi: 10.3758/s13428-011-0165

Forsythe, A., Mulhern, G., & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods, 40*(1), 116-129. doi: 10.3758/brm.40.1.116

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80*(1), 75-85. doi: 10.1037/0022-3514.80.1.75

Fussell, N. J., Rowe, A. C., & Mohr, C. (2012). Hemispheric processing of differently valenced and self-relevant attachment words in middle-aged married and separated individuals. *Laterality: Asymmetries of Body, Brain and Cognition, 17*(4), 453-485.

Garcia-Garcia, M., Dominguez-Borras, J., SanMiguel, I., & Escera, C. (2008). Electrophysiological and behavioral evidence of gender differences in the modulation of distraction by the emotional context. *Biological Psychology, 79*(3), 307-316.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion, 14*(2), 251-262.

Gilet, A. L., Gruehn, D., Studer, J., & Labouvie-Vief, G. (2012). Valence, arousal, and imagery ratings for 835 French attributes by young, middle-aged, and older adults: The French Emotional Evaluation List (FEEL). *European Review of Applied Psychology-Revue Europeenne De Psychologie Appliquee, 62*(3), 173-181. doi: 10.1016/j.erap.2012.03.003

Glenberg, A. M., Webster, B. J., Mouilso, E., Havas, D., & Lindeman, L. M. (2009). Gender, emotion, and the embodiment of language comprehension. *Emotion Review, 1*(2), 151-161

Goddard, C. (1997). Contrastive semantics and cultural psychology: 'Surprise' in Malay and English. *Culture & Psychology, 3*(2), 153-181. doi: 10.1177/1354067x9700300204

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition & Emotion, 22*(6), 1094-1118. doi: 10.1080/02699930701626582

Gohier, B., Senior, C., Brittain, P., Lounes, N., El-Hage, W., Law, V., . . . Surguladze, S. (2013). Gender differences in the sensitivity to negative stimuli: Cross-modal affective priming study. *European Psychiatry, 28*(2), 74-80. doi: http://dx.doi.org/10.1016/j.eurpsy.2011.06.007

Gole, M., Koechel, A., Schaefer, A., & Schienle, A. (2012). Threat engagement, disengagement, and sensitivity bias in worry-prone individuals as measured by an emotional go/no-go Task. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(1), 532-539.

Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R., & Winocur, G. (2006). Age-related changes in brain activity across the adult lifespan. [Article]. *Journal of Cognitive Neuroscience, 18*(2), 227-241.

Gratz, K. L., & Gunderson, J. G. (2006). Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. *Behavior Therapy, 37*(1), 25-35.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41-54.

Gruhn, D., & Scheibe, S. (2008). Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age? *Behavior Research Methods, 40*, 512–521.

Gruhn, D., & Smith, J. (2008). Characteristics for 200 words rated by young and older adults: Age-Dependent Evaluations of German Adjectives (AGE). *Behavior Research Methods, 40*(4), 1088-1097. doi: http://dx.doi.org/10.3758/BRM.40.4.1088

Gul, A., & Humphreys, G. W. (2014). Cultural effects in emotion and gender recognition. *Asian Journal of Social Psychology, 17*(1), 70-80.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin, 85*(4), 845.

Hall, J. A., Carter, J. D., & Horgan, T. G. (2000). Gender differences in nonverbal communications of emotion. In A. H. Fischer (Ed.), *Gender and emotions: Social psychological perspectives* (pp. 97-117). New York: Cambridge University Press.

Hall, J. A., & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion, 4*(2), 201-206.

Herrmann-Lingen, C., Buss, U., & Snaith, R. P. (2011). *Hospital anxiety and depression scale - German version (HADS-D)*. Bern: Hans Huber.

Hildebrandt, A., Wilhelm, O., Schmiedek, F., Herzmann, G., & Sommer, W. (2011). On the Specificity of Face Cognition Compared With General Cognitive Functioning Across Adult Age. *Psychology and Aging, 26*(3), 701-715. doi: 10.1037/a0023056

Hills, P. J., & Lewis, M. B. (2011). Reducing the own-race bias in face recognition by attentional shift using fixation crosses preceding the lower half of a face. *Visual Cognition, 19*(3), 313-339.

Hofer, A., Siedentopf, C. M., Ischebeck, A., Rettenbacher, M. A., Verius, M., Felber, S., & Fleischhacker, W. (2007). Sex differences in brain activation patterns during processing of positively and negatively valenced emotional words. *Psychological Medicine, 37*(1), 109-119.

Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S., & Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica, 135*(3), 278-283.

Horning, S. M. (2012). The recognition of facial expressions: An investigation of the influence of age, sex, and cognition. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 72*(7-B), 4320.

Hung, Y. (2014). Functional development of amygdalae and anterior cingulate cortex in emotion processing. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74*(8-B(E)), No Pagination Specified.

Hupka, R. B., Lenton, A. P., & Hutchison, K. A. (1999). Universal development of emotion categories in natural language. *Journal of Personality and Social Psychology, 77*(2), 247-278. doi: http://dx.doi.org/10.1037/0022-3514.77.2.247

Hutchison, A. N., & Gerstein, L. H. (2012). What's in a face? Counseling trainees' ability to read emotions. *Training and Education in Professional Psychology, 6*(2), 100-112.

Hyona, J., & Haikio, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology, 46*(6), 475-483.

Isaacowitz, D. M., & Stanley, J. T. (2011). Bringing an ecological perspective to the study of aging and recognition of emotional facial expressions: Past, current, and future methods. *Journal of Nonverbal Behavior, 35*(4), 261-278.

Ishii, K., Reyes, J. A., & Kitayama, S. (2003). Spontaneous attention to word content versus emotional tone: Differences among three cultures. *Psychological Science, 14*(1), 39-46.

Ito, K., Masuda, T., & Man Wai Li, L. (2013). Agency and facial emotion judgment in context. *Personality and Social Psychology Bulletin, 39*(6), 763-776.

Izard, C. E. (1990). Facial expressions and the regulation of emotions. *Journal of Personality and Social Psychology, 58*(3), 487-498.

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 109*(19), 7241-7244.

Juen, F., Huber, E. B., & Peham, D. (2012). Gender and age differences in emotion recognition of children and adolescent: First analysis of FACS coded childrens' faces. *Zeitschrift fur Entwicklungspsychologie und Pädagogische Psychologie, 44*(4), 178-191.

Kadosh, K. C., Johnson, M. H., Dick, F., Kadosh, R. C., & Blakemore, S.-J. (2013). Effects of age, task performance, and structural brain development on face processing. *Cerebral Cortex, 23*(7), 1630-1642.

Kanske, P., & Kotz, S. A. (2010). Leipzig Affective Norms for German: A reliability study. *Behavior Research Methods, 42*(4), 987-991. doi: 10.3758/brm.42.4.987

Kanske, P., & Kotz, S. A. (2011). Cross-modal validation of the Leipzig Affective Norms for German (LANG). *Behavior Research Methods, 43*(2), 409-413. doi: 10.3758/s13428-010-0048-6

Kayyal, M. H., & Russell, J. A. (2013a). Americans and Palestinians judge spontaneous facial expressions of emotion. *Emotion, 13*(5), 891-904.

Kayyal, M. H., & Russell, J. A. (2013b). Language and Emotion Certain English–Arabic Translations Are Not Equivalent. *Journal of Language and Social Psychology, 32*(3), 261-271.

Kellough, J. L., & Knight, B. G. (2012). Positivity effects in older adults' perception of facial emotion: The role of future time perspective. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 67B*(2), 150-158.

Kensinger, E. A. (2008). Age differences in memory for arousing and nonarousing emotional words. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 63B*(1), P13-P18. doi: http://dx.doi.org/10.1093/geronb/63.1.P13

Kinoshita, A., Okamoto, Y., Okada, G., Demoto, Y., Kunisato, Y., Yoshimura, S., et al. (2012). Sex differences in neural activation to ambiguous facial expression in happy and sad context. *Perceptual and Motor Skills, 115*(2), 349-359.

Knyazev, G. G., Slobodskoj-Plusnin, J. Y., & Bocharov, A. V. (2010). Gender Differences in Implicit and Explicit Processing of Emotional Facial Expressions as Revealed by Event-Related Theta Synchronization. *Emotion, 10*(5), 678-687.

Ko, S.-G., Lee, T.-H., Yoon, H.-Y., Kwon, J.-H., & Mather, M. (2011). How does context affect assessments of facial emotion? The role of culture and age. *Psychology and Aging, 26*(1), 48-59.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2013). Effects of aging on face identification and holistic face processing. *Vision Research, 88,* 38-46.

Lahl, O., Goeritz, A. S., Pietrowsky, R., & Rosenberg, J. (2009). Using the World-Wide Web to obtain large-scale word norms: 190,212 ratings on a set of 2,654 German nouns. *Behavior Research Methods, 41*(1), 13-19. doi: 10.3758/brm.41.1.13

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion, 12*(3), 529-539.

Landis, T. (2006). Emotional words: What's so different from just words? *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 42*(6), 823-830.

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson & T. A. Williams (Eds.), *Technology in mental health and delivery systems* (pp. 119-137). Norwood, NJ: Ablex.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International Affective Picture System (IAPS): Technical manual and affective ratings*. Gainesville: University of Florida, Center for Research in Psychophysiology.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures - Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261-273. doi: 10.1111/j.1469-8986.1993.tb03352.x

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion, 24*(8), 1377-1388. doi: Pii 930020275 10.1080/02699930903485076

Leclerc, C. M., & Kensinger, E. A. (2011). Neural processing of emotional pictures and words: A comparison of young and older adults. *Developmental Neuropsychology, 36*(4), 519-538.

Lee, N. C., Krabbendam, L., White, T. P., Meeter, M., Banaschewski, T., Barker, G. J., et al. (2013). Do you see what I see? Sex differences in the discrimination of facial emotions during adolescence. *Emotion, 13*(6), 1030-1040

Leime, J. L., Neto, J. R., Alves, S. M., & Torro-Alves, N. (2013). Recognition of facial expressions in children, young adults and elderly people. *Estudos de Psicologia, 30*(2), 161-167.

Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: How age shapes emotion processing in nonverbal vocalizations. *Emotion, 14*(1), 145-160.

Lundqvist, D., Flykt, A., & Oehman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Lundqvist, D., Svard, J., & Fischer, H. (2013). Age-related differences in sensitivity to emotional facial stimuli but age-independent association between arousal ratings and visual search efficiency. *Psihologijske Teme, 22*(2), 271-286.

Ma, Z., Li, J., Niu, Y., Yu, J., & Yang, L. (2013). Age differences in emotion recognition between Chinese younger and older adults. *The Psychological Record, 63*(3), 629-640.

Macchi Cassia, V., Luo, L., Pisacane, A., Li, H., & Lee, K. (2014). How race and age experiences shape young children's face processing abilities. *Journal of Experimental Child Psychology, 120*, 87-101.

Mammarella, N., Borella, E., Carretti, B., Leonardi, G., & Fairfield, B. (2013). Examining an emotion enhancement effect in working memory: Evidence from age-related differences. *Neuropsychological Rehabilitation, 23*(3), 416-428.

Mancini, G., Agnoli, S., Baldaro, B., Bitti, P. E. R., & Surcinelli, P. (2013). Facial Expressions of Emotions: Recognition Accuracy and Affective Reactions During Late Childhood. *Journal of Psychology, 147*(6), 599-617.

Marosi, E., Rivas, O., Yanez, G., & Bernal, J. (2002). Characterizing the emotional lexicon in Mexico in comparison with other cultures. *Revista Latina de Pensamiento y Lenguaje, 11*(1), 35-54.

Mathews, A., & Macleod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy, 23*(5), 563-569. doi: 10.1016/0005-7967(85)90104-4

Matsumoto, D., Olide, A., & Willingham, B. (2009). Is there an ingroup advantage in recognizing spontaneously expressed emotions? *Journal of Nonverbal Behavior, 33*(3), 181-191.

Mayr, U., Awh, E., & Laurey, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience, 6*(5), 450-452. doi: 10.1038/nn1051

McKenna, F. P. (1986). Effects of unattended emotional stimuli on color-naming performance. *Current Psychological Research & Reviews, 5*(1), 3-9. doi: http://dx.doi.org/10.1007/BF02686591

Melchers, M., Montag, C., Markett, S., & Reuter, M. (2013). Relationship between oxytocin receptor genotype and recognition of facial emotion. *Behavioral Neuroscience, 127*(5), 780-787.

Miellet, S., Vizioli, L., He, L., Zhou, X., & Caldara, R. (2013). Mapping face recognition information use across cultures. *Frontiers in Psychology, 4*, 34.

Mienaltowski, A., Johnson, E. R., Wittman, R., Wilson, A.-T., Sturycz, C., & Norman, J. (2013). The visual discrimination of negative facial expressions by younger and older adults. *Vision Research, 81*, 12-17.

Molnar, M., Toth, B., Boha, R., Gaal, Z., Kardos, Z., File, B., & Stam, C. (2013). Aging effects on erp correlates of emotional word discrimination. *Clinical Neurophysiology, 124*, 1986–1994. doi: http://dx.doi.org/10.1016/j.clinph.2013.04.017

Moore, C. C., Romney, A. K., Hsia, T. L., & Rusch, C. D. (1999). The universality of the semantic structure of emotion terms: Methods for the study of inter- and intra-cultural variability. *American Anthropologist, 101*(3), 529-546. doi: 10.1525/aa.1999.101.3.529

Murray, B. D., & Kensinger, E. A. (2013). Age-related changes in associative memory for emotional and nonemotional integrative representations. *Psychology and Aging, 28*(4), 969-983. doi: http://dx.doi.org/10.1037/a0034443

Noh, S. R., & Isaacowitz, D. M. (2013). Emotional faces in context: Age differences in recognition accuracy and scanning patterns. *Emotion, 13*(2), 238-249.

Osgood, C., Suci, G., & Tannenbaum, P. (1957). The measurement of meaning. Urbana, IL: University of Illinois.

Pagliaccio, D., Luby, J. L., Gaffrey, M. S., Belden, A. C., Botteron, K. N., Harms, M. P., et al. (2013). Functional brain activation to emotional and nonemotional faces in healthy children: Evidence for developmentally undifferentiated amygdala function during the school-age period. *Cognitive, Affective & Behavioral Neuroscience, 13*(4), 771-789.

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods Instruments & Computers, 36*(4), 634-638. doi: 10.3758/bf03206544

Panayiotou, A. (2002). The other within the self: Bilinguals and the construction of emotions. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 62*(11-B), 5431.

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism-Language and Cognition, 11*(2), 147-164

Prado, C., Mellor, D., Byrne, L. K., Wilson, C., Xu, X., & Liu, H. (2013). Facial emotion recognition: A cross-cultural comparison of chinese, chinese living in australia, and anglo-australians. *Motivation and Emotion Nov*(Pagination), No Pagination Specified.

Rahman, Q., & Anchassi, T. (2012). Men appear more lateralized when noticing emotion in male faces. *Emotion, 12*(1), 174-179.

Redondo, J., Fraga, I., Padron, I., & Comesna, M. (2007). The Spanish adaptation of ANEW (Affective Norms for English words). *Behavior Research Methods, 39*(3), 600-605. doi: 10.3758/bf03193031

Rennels, J. L., & Cummings, A. J. (2013). Sex differences in facial scanning: Similarities and dissimilarities between infants and adults. *International Journal of Behavioral Development, 37*(2), 111-117.

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin, 138*(1), 146-174.

Riediger, M., Voelkle, M. C., Ebner, N. C., & Lindenberger, U. (2011). Beyond "happy, angry, or sad?": Age-of-poser and age-of-rater effects on multi-dimensional emotion perception. *Cognition and Emotion, 25*(6), 968-982.

Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of Facial, Auditory, and Bodily Emotions in Older Adults. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 64*(6), 696-703.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews, 32*(4), 863-881.

Rusch, C. D. (1997). The effects of bilingualism on cognitive semantic structure in the subjective lexicon: The case of emotions in Japanese and English. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 57*(10-A), 4428.

Rusch, C. D. (2004). Cross-cultural variability of the semantic domain of emotion terms: An examination of English shame and embarrass with Japanese hazukashii. *Cross-Cultural Research: The Journal of Comparative Social Science, 38*(3), 236-248.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression - A review of the cross-cultural studies. *Psychological Bulletin, 115*(1), 102-141.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145-172. doi: 10.1037/0033-295x.110.1.145

Russell, J. A., & Sato, K. (1995). Comparing emotion words between languages. *Journal of Cross-Cultural Psychology, 26*(4), 384-391.

Sasson, N. J., Pinkham, A. E., Richard, J., Hughett, P., Gur, R. E., & Gur, R. C. (2010). Controlling for response biases clarifies sex and age differences in facial affect recognition. *Journal of Nonverbal Behavior, 34*(4), 207-221. doi: http://dx.doi.org/10.1007/s10919-010-0092-z

Schmidt-Atzert, L., & Park, H.-S. (1999). The Korean concepts dapdaphada and uulhada: A cross-cultural study of the meaning of emotions. *Journal of Cross-Cultural Psychology, 30*(5), 646-654.

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology, 46*(6), 401-435. doi: 10.1080/00207594.2011.626049

Schwabe, L., Hoffken, O., Tegenthoff, M., & Wolf, O. T. (2013). Opposite effects of noradrenergic arousal on amygdala processing of fearful faces in men and women. *Neuroimage, 73*, 1-7.

Senholzi, K. B., & Ito, T. A. (2013). Structural face encoding: How task affects the N170's sensitivity to race. *Social Cognitive and Affective Neuroscience, 8*(8), 937-942.

Silk, J. S., Siegle, G. J., Whalen, D. J., Ostapenko, L. J., Ladouceur, C. D., & Dahl, R. E. (2009). Pubertal changes in emotional information processing: Pupillary, behavioral, and subjective evidence during emotional word identification. *Development and Psychopathology, 21*(1), 7-26.

Soares, A. P., Comesana, M., Pinheiro, A. P., Simoes, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. doi: 10.3758/s13428-011-0131-7

Soederholm, C., Hayry, E., Laine, M., & Karrasch, M. (2013). Valence and Arousal Ratings for 420 Finnish Nouns by Age and Gender. *Plos One, 8*(8). doi: e7285910.1371/journal.pone.0072859

Sommer, W., Hildebrandt, A., Kunina-Habenicht, O., Schacht, A., & Wilhelm, O. (2013). Sex differences in face cognition. *Acta Psychologica, 142*(1), 62-73.

Stanley, J. T., Zhang, X., Fung, H. H., & Isaacowitz, D. M. (2013). Cultural differences in gaze and emotion recognition: Americans contrast more than Chinese. *Emotion, 13*(1), 36-46.

Stevenson, R. A., Mikels, J. A., & James, T. W. (2007). Characterization of the affective norms for english words by discrete emotional categories. *Behavior Research Methods, 39*(4), 1020-1024. doi: 10.3758/bf03192999

Stoesz, B. M., & Jakobson, L. S. (2013). A sex difference in interference between identity and expression judgments with static but not dynamic faces. *Journal of Vision, 13*(5), 26.

Strick, M., Holland, R. W., & van Knippenberg, A. (2008). Seductive eyes: Attractiveness and direct gaze increase desire for associated objects. *Cognition, 106*(3), 1487-1496. doi: 10.1016/j.cognition.2007.05.008

Suzuki, A., & Akiyama, H. (2013). Cognitive aging explains age-related differences in face-based recognition of basic emotions except for anger and disgust. *Aging, Neuropsychology, and Cognition, 20*(3), 253-270.

Syssau, A., & Monnier, C. (2009). Children's emotional norms for 600 French words. *Behavior Research Methods, 41*(1), 213-219.

Tahmasebi, A. M., Artiges, E., Banaschewski, T., Barker, G. J., Bruehl, R., Buchel, C., et al. (2012). Creating probabilistic maps of the face network in the adolescent brain: A multicentre functional MRI study. *Human Brain Mapping, 33*(4), 938-957.

Thapar, A., & Rouder, J. N. (2009). Aging and recognition memory for emotional words: A bias account. *Psychonomic Bulletin & Review, 16*(4), 699-704. doi: http://dx.doi.org/10.3758/PBR.16.4.699

Thayer, J. F., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology, 41*(3), 243-246.

Thibault, P., Levesque, M., Gosselin, P., & Hess, U. (2012). The Duchenne marker is not a universal signal of smile authenticity-But it can be learned! *Social Psychology, 43*(4), 215-221. doi: http://dx.doi.org/10.1027/1864-9335/a000122

Thomas, R. C., & Hasher, L. (2006). The influence of emotional valence on age differences in early processing and memory. *Psychology and Aging, 21*(4), 821-825. doi: http://dx.doi.org/10.1037/0882-7974.21.4.821

Torro-Alves, N., de Oliveira Bezerra, I. A., Claudino, R. G. e., & Pereira Cavalcanti, T. L. (2013). Influences of sex, type and intensity of emotion in the ecognition of static and dynamic facial expressions. *Avances en Psicologia Latinoamericana, 31*(1), 192-199.

Tortosa, M. I., Lupianez, J., & Ruz, M. (2013). Race, emotion and trust: An ERP study. *Brain Research, 1494*, 44-55.

Tottenham, N., Phuong, J., Flannery, J., Gabard-Durnam, L., & Goff, B. (2013). A negativity bias for ambiguous facial-expression valence during childhood: Converging evidence from behavior and facial corrugator muscle responses. *Emotion, 13*(1), 92-103.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242-249.

Tracy, J. L., Robins, R. W., & Schriber, R. A. (2009). Development of a FACS-Verified Set of Basic and Self-Conscious Emotion Expressions. *Emotion, 9*(4), 554-559. doi: 10.1037/a0015766

Triandis, H. C., & Osgood, C. E. (1958). A comparative factorial analysis of semantic structures in monolingual Greek and American college students. *Journal of Abnormal and Social Psychology, 57*(2), 187-196.

Voe, M. L. H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL–R). *Behavior Research Methods, 41*(2), 534-538.

Voe, M. L. H., Jacobs, A. M., & Conrad, M. (2006). Cross-validating the Berlin Affective Word List. *Behavior Research Methods, 38*(4), 606-609.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental Science, 15*(3), 359-372.

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior, 17*(1), 3-28. doi: 10.1007/bf00987006

Wallis, G. (2013). Toward a unified model of face and object recognition in the human visual system. *Frontiers in Psychology, 4*, 497.

Wang, B. (2013). Gender difference in recognition memory for neutral and emotional faces. *Memory, 21*(8), 991-1003.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191-1207. doi: 10.3758/s13428-012-0314-x

Weidner, K. E. (2014). Recognizing emotional faces: The role of gender, age, and attractiveness. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 74*(8-B(E)), No Pagination Specified.

West, J. T., Horning, S. M., Klebe, K. J., Foster, S. M., Cornwell, R., Perrett, D., et al. (2012). Age effects on emotion recognition in facial displays: From 20 to 89 years of age. *Experimental Aging Research, 38*(2), 146-168.

Wickline, V. B., Bailey, W., & Nowicki, S. (2009). Cultural in-group advantage: Emotion recognition in African American and European American faces and voices. *The Journal of Genetic Psychology: Research and Theory on Human Development, 170*(1), 5-29. doi: http://dx.doi.org/10.3200/GNTP.170.1.5-30

Wierzbicka, A. (2010). Cross-cultural communication and miscommunication: The role of cultural keywords. *Intercultural Pragmatics, 7*(1), 1-23. doi: 10.1515/iprg.2010.001

Willinger, U., Schmoeger, M., Deckert, M., & Auff, E. (2012). Viennese Emotional Go/no-go Task (V–EGT). Unpublished computerized test.

Willinger, U., Schmoeger, M., Deckert, M., & Auff, E. (2012). Viennese Emotional Stroop Task (V-EST). Unpublished computerized test.

Willinger, U., Schmoeger, M., Deckert, M., & Auff, E. (2013).Vienese Affective Flexibility Task (V-AFT). Unpublished computerized test.

Wurm, L. H., Labouvie-Vief, G., Aycock, J., Rebucal, K. A., & Koch, H. E. (2004). Performance in Auditory and Visual Emotional Stroop Tasks: A Comparison of Older and Younger Adults. *Psychology and Aging, 19*(3), 523-535. doi: http://dx.doi.org/10.1037/0882-7974.19.3.523

Xu, Q., Yang, Y., Wang, P., Sun, G., & Zhao, L. (2013). Gender differences in preattentive processing of facial expressions: An ERP study. *Brain Topography, 26*(3), 488-500.

Yang, L., & Hasher, L. (2011). Age differences in the automatic accessibility of emotional words from semantic memory. *Cognition and Emotion, 25*(1), 3-9.

Zhu, Z., Ho, S. M., & Bonanno, G. A. (2013). Cultural similarities and differences in the of emotional valence and intensity: A comparison of Americans and Hong Kong Chinese. *The American Journal of Psychology, 126*(3), 261-273.

List of tables and figures

List of tables:

List of figures:

Appendix

A1: Demographic data form

Soziodemographischer Fragebogen

Geschlecht: O Weiblich O Männlich

Alter: ____ Jahre

Herkunft:
In welchem Land und welchem Bundesland:

- sind Sie **geboren**?
  Land:_____ Bundesland: _____
- haben Sie den Großteil Ihrer **Kindheit** verbracht?
  Land:_____ Bundesland: _____
- haben Sie den Großteil Ihrer **Jugend** verbracht?
  Land:_____ Bundesland: _____
- haben Sie den Großteil Ihres **Erwachsenenalters** verbracht?
- Land:_____ Bundesland: _____

Sprache:
Welche Sprache ist Ihre Muttersprache? _____

Welche Sprachen sprechen Sie fließend und wie oft wenden Sie diese an?

_____ O Sehr selten O Selten O Manchmal O Häufig O Sehr häufig
_____ O Sehr selten O Selten O Manchmal O Häufig O Sehr häufig
_____ O Sehr selten O Selten O Manchmal O Häufig O Sehr häufig
_____ O Sehr selten O Selten O Manchmal O Häufig O Sehr häufig
_____ O Sehr selten O Selten O Manchmal O Häufig O Sehr häufig

Händigkeit: O Links O Rechts O Beidhändig

Religion: Sind Sie einer Religion zugehörig? O Nein O Ja - Wenn Ja, welche: _____

Höchste abgeschlossene Ausbildung:

O Pflichtschule
O Fachschule
O Studienberechtigungsprüfung
O Matura bzw. Abitur
O Akademie
O Fachhochschule
O Universität
O Sonstige: _____

Akademischer Grad: _____

Welchen Schultyp haben Sie besucht:_____

Bitte umblättern! →

1

Absolvieren Sie zurzeit ein Universitätsstudium?
O nein O ja
Studienrichtung: _____
Im wievielten Semester befinden Sie sich? _____

Berufstätigkeit: O nein O ja Wenn ja, wie viele Stunden pro Woche? _____

Ausgeübter Beruf:_____

Erkrankungen bzw. Beeinträchtigungen:
Sind Sie kurzsichtig? O nein O ja
Sind Sie weitsichtig? O nein O ja
Wenn Ja, gleichen Sie diese zurzeit durch eine Brille bzw. Kontaktlinsen aus? O nein O ja

Haben Sie eine eingeschränkte Sehfähigkeit bzw. Erkrankungen, die eine beeinträchtigte Wahrnehmung bewirken (ausgenommen Kurz- und Weitsichtigkeit), wie z.B. Farbenblindheit?

O Nein O Ja
Wenn ja welche:_____

Leiden oder litten Sie an neurologischen oder psychiatrischen Erkrankungen?

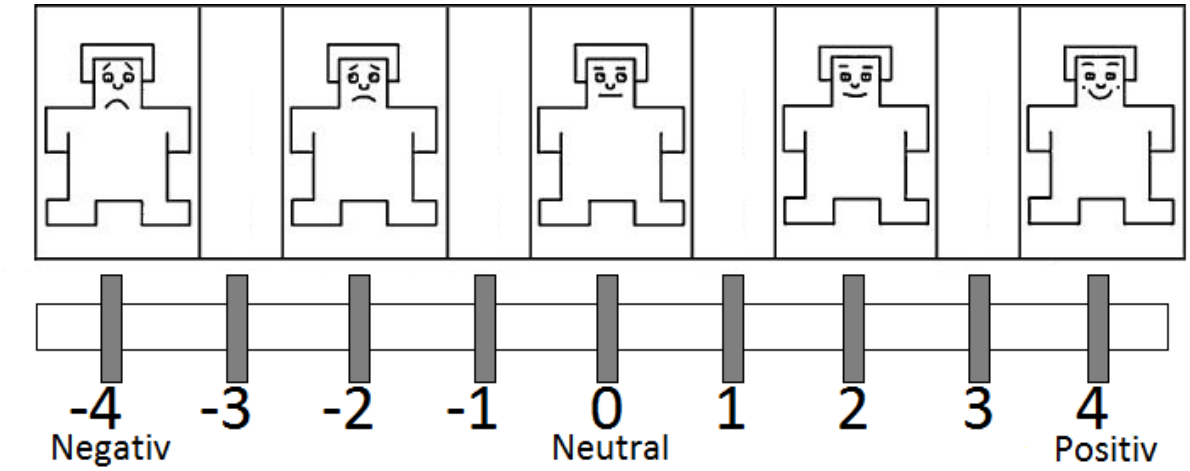
O Nein O Ja
Wenn ja welche:_____
_____
_____
_____

Welche von ihnen sind akut:
_____
_____
_____
_____

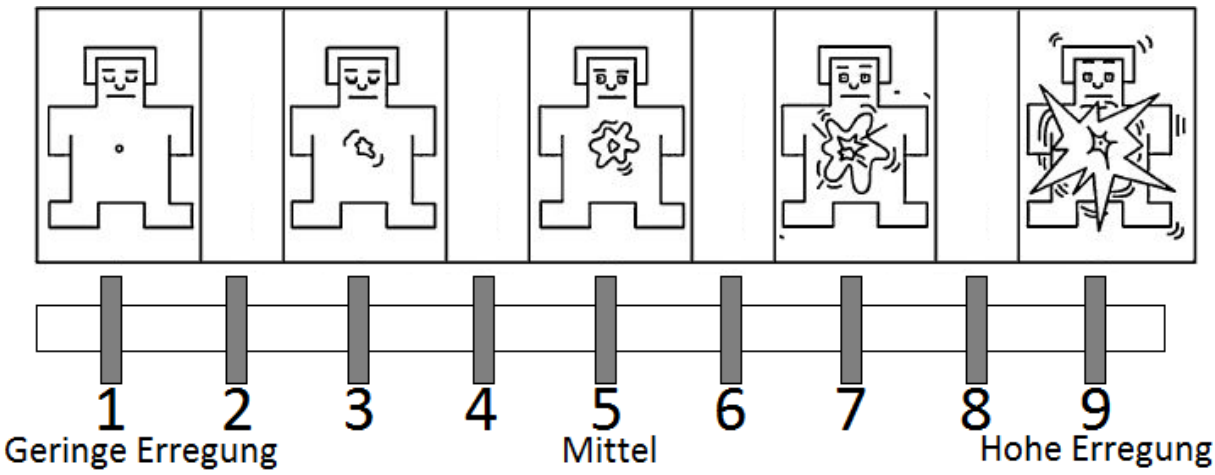
Medikamente:
Nehmen Sie zurzeit Medikamente? O Nein O Ja

Wenn Ja, welche nehmen Sie zurzeit? _____
_____
_____
_____
_____
_____

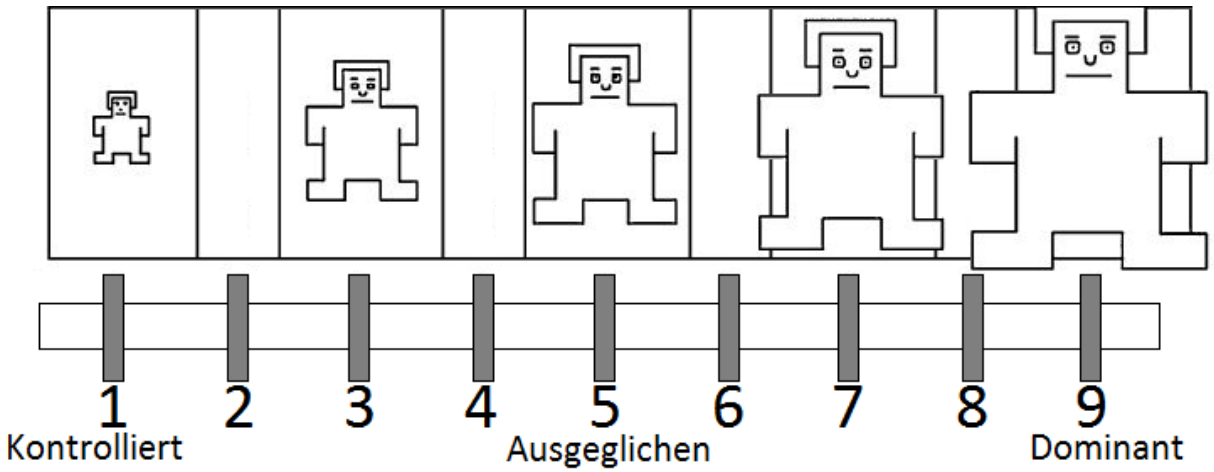
A2: Valence rating scale (German) - Bradley & Lang (1994)



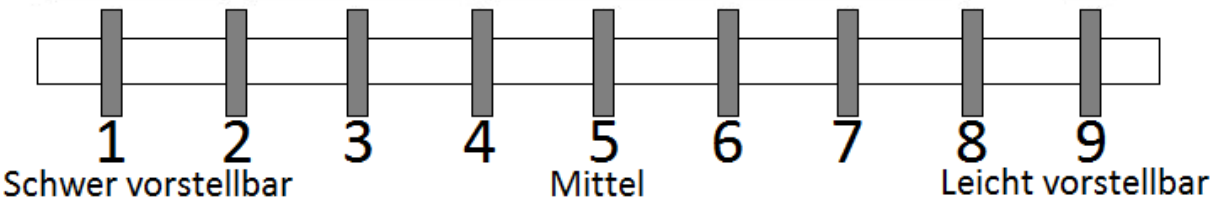
A3: Arousal rating scale (German) - Bradley & Lang (1994)



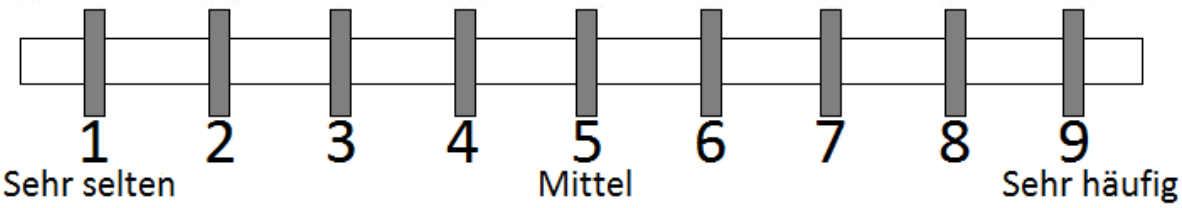
A4: Dominance rating scale (German) - Bradley & Lang (1994)



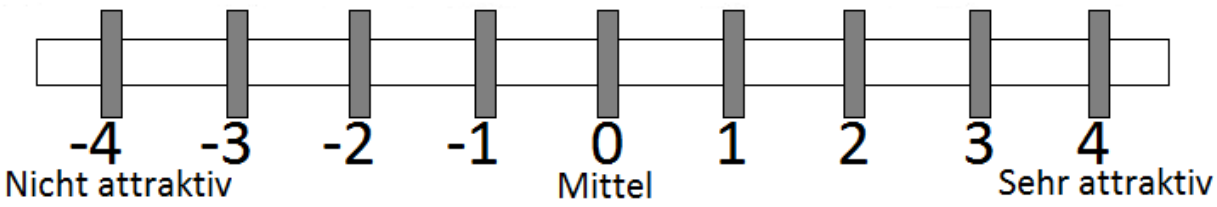
A5: Imagery rating scale (German)



A6: Familiarity rating scale (German)



A7: Attractiveness scale (German)



A8: Emotion recognition I (German) – Picture taken from the KDEF (Lundqvist et al. 1998)

Emotion:
Welche Emotion wird hier dargestellt?

Freude	Angst	Trauer	Ärger	Neutral	Keine der Genannten

A9: Emotion recognition II (German) - Picture taken from the KDEF (Lundqvist et al. 1998)

Emotion:
Welche Emotion wird hier dargestellt?

Freude

Angst

Curriculum Vitae

Schulbildung:

1994-1998:	Volksschule 1100 Wien, Oberlaaer Platz 1
1998-2006:	Bundesrealgymnasium 1100 Wien, Pichelmayergasse 1
Mai/Juni 2006:	Matura am Bundesrealgymnasium 1100 Wien, Pichelmayergasse 1
Oktober 2006 – Juni 2007	Ausbildung zum Rettungssanitäter und Zivildienst beim Österreichischen Roten Kreuz
2007-2009:	Bachelorstudium Violoncello an der Bruckner Privatuniversität 4040 Linz
Ab 2008:	Ordentliches Psychologiestudium in Wien
Dezember 2010:	1.Diplomprüfung Psychologie absolviert
Februar 2014:	Abschluss 1.Teil der 2. Diplomprüfung Psychologie

Berufstätigkeit/Praktika:

Juli 2012- Dezember 2012	Praktikum an der Universitätsklinik für Neurologie, Medizinische Universität Wien
Ab Dezember 2012	Tätigkeit als nicht-wissenschaftlicher Mitarbeiter an der Universitätsklinik für Neurologie, Medizinische Universität Wien

Fremdsprachen:

Englisch	(Matura)
Isländisch	(Zweisprachig von Geburt an)