Is Emotion more strongly affected by Gender or Motion?

On the influence of the pianist's biological gender on the perception of gesturally expressed emotions in musical performance: An empirical case study.

A thesis submitted in partial fulfilment of the requirements for the degree of

MASTER OF ARTS (MA)

of

UNIVERSITY of GRAZ
(Department of Musicology)
&
UNIVERSITY of MUSIC and PERFORMING ARTS GRAZ
(Institute of Electronic Music and Acoustics)

by

JOHANNES LEHNER

Graz, 2013

Thesis Advisor: Univ.-Prof. Dr. Gerhard Eckel

KARL-FRANZENS-UNIVERSITÄT GRAZ
UNIVERSITY OF GRAZ

Abstract

Prejudices and social norms often affect our value judgements – especially with respect to gender stereotypes. Emotions are mainly communicated via body movement. Based on these suppositions and sociological and music-psychological studies, this thesis explores (biological and social) gender-specific perception of gesturally expressed emotions in musical performance of pianists. The aim of this work was to investigate the influence of the biological gender of musicians on the perception of intensity of four emotion categories (anger, fear, happiness, and sadness). For this purpose an experiment has been conducted with 32 students. A total of 9 different silent videos of pianists playing were presented to the participants: The videos were motion capture recordings of 3 different pianists' upper bodies playing in 3 different expression levels. All videos were played 5 times, labelled twice with a female and a male name, and once with a gender neutral name. The hypothesised impact of the labels couldn't be validated, instead, the results have shown with high statistical significance that expressive motion affects our perception of emotions much more than the labelling. Moreover, on the evidence presented, no clear conclusion can be drawn regarding the sub-theses that the rating of emotion depends on the pianists' biological gender and/or the participants' biological or social gender.

Keywords:
Perception of Emotion • Gender Differences • Biological Gender • Social Gender • Body Movement • Motion Capture

Kurzfassung

Vorurteile und gesellschaftliche Normen bestimmen oftmals unsere Werturteile – speziell in Bezug auf Geschlechterstereotypen. Die Kommunikation von Emotionen erfolgt hauptsächlich über Körpersprache. Ausgehend von diesen Hyothesen und soziologischen und musikpsychologischen Studien, beschäftigt sich diese Arbeit mit geschlechter- und gender-spezifischer Wahrnehmung von gestisch vermittelten Emotionen in musikalischen Aufführungen von PianistInnen. Ziel war die Untersuchung des Einflusses des biologischen Geschlechts der MusikerInnen auf die Wahrnehmung der Intensität von vier Emotionskategorien (Zorn, Angst, Freude, und Trauer). Dazu wurde ein Experiment mit 32 Studierenden durchgeführt. Insgesamt wurden den VersuchsteilnehmerInnen 9 verschiedene stumme Videos von PianistInnen während des Spielens vorgeführt: Bei den Videos handelte es sich um Motion Capture Aufnahmen der Oberkörper von 3 unterschiedlichen PianistInnen in jeweils 3 Ausdrucksstärken. Jedes Video wurde insgesamt 5 Mal vorgespielt und wurde je zweimal mit einem *männlichen* und *weiblichen* Namen, sowie einmal mit einem *geschlechtsneutralen* Namen gekennzeichnet. Der vermutete Einfluss dieser Geschlechtszuschreibung konnte nicht bestätigt werden, vielmehr haben die Ergebnisse mit hoher statistischer Signifikanz gezeigt, dass ausdrucksstarke Bewegung die Wahrnehmung von Emotionen viel stärker beeinflusst als die Geschlechtszuschreibung über Namen. Aufgrund der vorliegenden Ergebnisse kann außerdem keine klare Schlussfolgerung in Bezug auf die Sub-Thesen, dass die Bewertung von wahrgenommenen Emotionen von dem Geschlecht des/der PianistIn und/oder dem biologischen oder sozialen Geschlecht des/der ProbandIn abhängt, gezogen werden.

Stichworte:
Emotionswahrnehmung • Geschlechtsunterschiede • Biologisches Geschlecht • Soziales Geschlecht • (körperliche) Bewegung • Motion Capture

Acknowledgements

This master thesis has been carried out at the Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz, in cooperation with the Department of Musicology, University of Graz. A number of people deserve thanks for their support and help. It is therefore my greatest pleasure to express my gratitude to them all in this acknowledgement.

First and foremost I would like to express my deep gratitude to my master thesis advisor Dr. Gerhard Eckel, who has supported me throughout my thesis. He is also to be thanked for introducing me into PureData.

Secondly, it is a pleasure to thank Marc Thompson for kindly providing his motion capture video recordings, and Dr. Ursula Athenstaedt for providing the German Extended Personal Attributes Questionnaire for this study without which the experiment would not have been possible to this extent. In addition, I thank Peter Gaube for supporting me in sociological questions.

Furthermore, I would also like to acknowledge with much appreciation the staff of the Institute of Electronic Music and Acoustics, who has provided the support and equipment as well as the location I have needed to conduct my experiment. In particular, I would like to thank DI Johannes Zmölnig for his generous advice and support in technical questions concerning PureData.

A special thanks goes to my girlfriend, Nadine Scharfetter, who assisted me as a second investigator in performing the main experiment. Also, many thanks go to Priscilla Aistleitner and David Roode for proof-reading my thesis.

Last but not least, I would like to thank the participants volunteered in my two pre-experiments, who have willingly shared their precious time, particularly Christoph Nausner, who has agreed to test both pre-experiments before carrying them out.

Graz, August 2013 Johannes Lehner

Table of Contents

List of Tables

List of Figures

1 Introduction

In recent years, many studies have focused on (communication of) emotion (for examples, see Ekman, 2003, Juslin & Laukka, 2003, Hall & Matsumoto, 2004, Brody & Hall, 2008), research about body movement and motion (e.g., Dittrich et al., 1996, Wallbott, 1998, Clarke et al., 2005), as well as biological and social gender issues[1] (for an overview, see Kreutziger-Herr & Unseld, 2010, Ellis, 2008). Previous research has concentrated on considering the issues separately, or in combination of only two of them (examples: Pollick et al., 2001, Fischer et al., 2004, Clarke et al., 2005, Dahl & Friberg, 2007), and comparatively few attempts have, so far, been made to link all three topics so far (especially in musicological research). Therefore, this paper aims to link those three fields together and consider them as one single issue embedded in a musicological context.

Considering all aspects of the three fields would go far beyond the scope of this paper, therefore this master thesis focuses on 4 aspects: The influence of the (1) artists' biological gender on the perception of (2) gesturally expressed (3) (basic) emotions (4) in music. Backgrounds to each of these aspects are briefly summarised in the following paragraphs:

Overview

Emotion. Research on expression of emotions and their meaning has a long history. Darwin (1872/1965) was the first[2] to document facial expressions as an important source for decoding emotional cues. Shariff and Tracy (2011) refer in their article to the evolutionary origins and functions of emotional expressions, and discuss the importance of nonverbal expression of emotion in communication of emotions. This role might be the reason for the great number of research in this field (e.g. Walk & Homan, 1984, Hall & Matsumoto, 2004, Gosselin et al., 2005, Meeren et al., 2005).

1 Biological and social gender issues are sometimes closely connected in research, but in some cases they differ widely from each other. Gender differences are culturally influenced in most cases and therefore a linking issue of both fields.

2 No earlier sources were found by the author. Several other publications (e.g., Ekman, 1999, Meeren et al., 2005, Avizier, 2008) support this statement.

Drawing on previous studies, two major (basic) emotion theories[3] have emerged (see Barrett, 1998, Thamm, 2006, Aviezer et al., 2008, Clore & Ortony, 2008, Niedenthal, 2008, Barrett, 2011): First, the discrete category view which is based on Darwin's theory of functional expressions of emotions. It assumes that a small amount of basic emotions[4] are evolutionary determined, and all sub-types of emotions like pride, optimism, relief, or disappointment (see Shaver et al., 1987) refer to those basic emotions. Secondly, the dimensional view explains emotions through a dimensional structure (for further reading, see Thamm, 2006, Niedenthal, 2008) as an alternative to the discrete view. Dimensional theories structure emotions in a continuum of two or more dimensions such as arousal, evaluation, potency, valence, activity, and/or pleasantness. According to Barrett (1998) and Brosch et al. (2010), both emotional theories can be appropriate but also have their limitations. Therefore, this study builds on earlier musicological research (e.g., Juslin & Laukka, 2003, Resnicow et al., 2004, Dahl & Friberg, 2007) and is set up using a categorical system.

Recent research (Meeren et al., 2005, Lindquist et al., 2006, Aviezer et al., 2008) has shown that the readout of specific emotional categories is affected by their context, no matter if the context is communicated via the same channel (e.g., Lindquist et al., 2006). On the contrary, several studies have shown that at least some emotional cues can be decoded even if parts of the body are hidden or only reduced information[5] is seen by the receiver (Walk & Homan, 1984, Dittrich et al., 1996, Dahl & Friberg, 2004, Clarke et al., 2005). However, those findings don't exclude each other completely. Moreover, it can be concluded that not all (sub-)categories of emotions use the same channel of communication. Even if the communication of emotion is assumed to be more closely linked with facial or vocal expression in research (e.g., Darwin, 1872/1965, Ekman, 2003, Juslin & Laukka, 2003, Gosselin et al., 2005), previous research has shown that basic emotions can be well communicated through gestures and body movement (see Wallbott, 1998, Montepare et al., 1999).

3 Based on those basic theories several sub-theories have been developed (for further reading, see Thamm, 2006, Niedenthal, 2008).

4 The basic emotions differ from theory to theory (see Ortony & Turner, 1990). Often mentioned basic emotions are anger, disgust, fear, happiness, sadness, and surprise.

5 Examples for reduced context/information are: Blurred faces, or recordings of persons reduced to point-lights.

In addition, several other aspects have been documented to affect the process of decoding emotional cues. Most important to mention are differences in culture[6] on the one hand, whereby they are reported not to be the same for all emotion categories (see Ekman, 1999, Shariff & Tracy, 2011), and on the other hand, (biological) gender differences in decoding nonverbal emotion cues (see Locke, 2002, Wester et al., 2002, Brody & Hall, 2008). In particular, this study considers two aspects in context with emotion and music: (1) Body movement, and (2) gender differences.

Motion & Emotion. Body movement, or motion, as it is specified in this paper, referring to the motion capture videos, is an important part of social interaction. Several studies already demonstrated that body movement is often used to express emotions. For example, Wallbott (1998) showed that some body movements are specific for certain basic emotions, building on Darwin's *'The Expression of Emotions in Man and Animals'* (1872/1965). Furthermore, a number of researchers (e.g. Walk & Homan, 1984, Dittrich et al., 1996, Montepare et al., 1999, Dahl & Friberg, 2004, Clarke et al., 2005) have demonstrated that we are still able to decode basic emotions very accurately even if the body movement is shown only by some light points. Despite small differences in their findings (e.g., in the accuracy of identifying fear), it has been shown that the emotions Happiness, Sadness, Anger, and Fear seem to be consistently communicated very accurately through motion (see Dahl & Friberg, 2004, 2007; Clarke et al., 2005).

Besides every-day movements such as, for example, walking or lifting the arm to drink, body movement has also been deeply linked to musical performance. Davidson and Correia (2002) argue that even if the production of sounds by body movement[7] is the most obvious link, it is not the only one. Moreover, they mention several aspects that influence motion and therefore emotional content in performances (e.g., the performer's past experiences, communication and interaction with co-artists and the

6 Shariff and Tracy (2011) report that from a biological point of view it is important to make a distinction between *cues* and *signals*. Cues provide information as a by-product and are basically not affected by culture, and signals evolved to communicate information to others but are culturally dependent.

7 Sounds from many typical acoustic instruments like guitar, piano, or drums are produced through body movement.

audience). From a biological point of view, the *basic* movements of artists manipulating their instruments could be seen as *cues* (see Shariff & Tracy, 2011), and all additional motion as *signals*. However, there is a large amount of research on dance and music performances (e.g., Walk & Homan, 1984; Dittrich et al., 1996; Resnicow et al., 2004; Dahl & Friberg, 2004, 2007) concerning issues affected by signals. Even if emotional content is more accurately communicated by signals, Tabei and Tanaka (2012) found that perception of emotion depends on the combination of the played instrument and emotion. Furthermore, Dahl and Friberg (2004) discussed findings by Sörgjerd in 2000 in their article: Sörgjerd's thesis showed that the emotions Happiness, Sadness, Anger, and Fear were decoded more accurately in body movement than other basic emotions. The same emotions were used throughout several other musicological studies (e.g., Juslin & Laukka, 2003, Resnicow et al., 2004, Tabei & Tanaka, 2012). Hence, the emotions Happiness, Sadness, Anger, and Fear might be best for a study linking the fields of body movement and of emotion research together.

Culture & Emotion. As mentioned earlier in this paper, cultural differences affect research on emotion although they are not always present. Matsumoto et al. (2002) suggested several reasons for differences in ratings of emotions, whereby semantic overlap and the culture itself might be the most important two. Locke (2002) reports in her review that "in terms of feminist thought, the [biological] gender differences that appear in emotionality can be seen as being due to cultural expectations of emotional expression [...]" (p.97). Furthermore, she discusses the stereotypes of the 'emotional female' and the 'non-emotional, rational male' as culturally coded values. Partial proof can be found in Fischer et al. (2004), who argues that some emotions are determined more strongly by biological factors, and others more by social roles (and therefore by culture)[8]. However, they have one important finding in common: The articles report gender differences as important aspect of cultural differences (see Locke, 2002, Matsumoto et al., 2002, Fischer et al., 2004, Brody & Hall, 2008).

8 Anger has been reported to be strongly influenced by the social status of women (see Fischer et al., 2004). In general, anger is seen as an emotion with great gender differences in expression and decoding abilities (see Milovchevich et al., 2001).

Biological and social Gender, & Emotion. Biological and social gender differences are not only an inter-cultural, but also an inner-cultural issue (see Hall et al., 2000, Locke, 2002, Fischer et al., 2004). In particular, a lot of research on decoding nonverbal cues and emotional expression has already been done in the field of biological and social gender studies (e.g., Hall, 1978, 1984; Hall & Matsumoto, 2004; Fischer et al., 2004; Rosip & Hall, 2004). Still, the findings differ from each other: According to Rosip and Hall (2004) and Ickes et al. (2000), females have a higher nonverbal decoding accuracy and a higher empathic accuracy than men. As Hall et al. (2000) have noted, "differences between males' and females' nonverbal behaviours and skills have evolved biologically" (p. 99), but on the other hand, they argued that female and male stereotypes are still affected by society. In addition, Wester et al. (2002) reported that (biological) gender differences in decoding accuracy of emotions seem to disappear in western society. Recent research (see Shields et al., 2006) discussed the aspects reported by Wester et al., and argued that gender isn't static like biological gender[9], but more something we are 'doing'[10]. In the same context Shields et al. also suggested that emotion is something we are doing and not something we have. In fact, the socialised difference in decoding accuracy skills is only a single aspect of biological and social gender issues in emotion research.

Besides the biological and social gender differences in decoding accuracy of emotions, the biological gender of actors and performers, or even people we talk to in every-day life affects our ratings: Several researchers (e.g., Hall, 1978, Behne, 1990, Locke, 2002, Brody & Hall, 2008) have reported that prejudices often affect how emotional or rational we rate and judge others. Summarising several studies, Hall (1978) validated the theory even if it was not consistent in all studies. In the same context, Behne (1990) discussed two important findings about ratings on piano performances in particular: (1) Male performers were rated to be more precise in playing, whilst female performers were seen as being more emotional, and (2) ratings about musical performances actually are value judgements about the look, the attractiveness, the sympathy, and the charisma of the performer.

9 This paper only refers to natural changes and not gender change using modern medicine.

10 Gender studies use the term 'Doing Gender' to explain the dynamic process/interaction that goes together with considering genders in a socio-cultural context.

Aims of the Study

Main Hypothesis. Primarly, this study aims to answer the question: Is the communication of emotions more strongly affected by the biological gender of artists or by their expressions in movement? Based on this question several assumptions were made: Considering the findings of Hall (1978), Behne (1990), Locke (2002), and Brody and Hall (2008), it was hypothesised that the communication of emotions is more strongly affected by the biological gender of the artist than it is by his or her bodily expression.

Sub-Theses. Moreover, it was assumed that the emotion category and the rated intensity of these emotions are affected by the performer's biological gender. Recent research (e.g., Locke, 2002, Fischer et al., 2004) discussed that the social status and culture affect females' and males' stereotypes in society. Therefore, it was assumed that female performers are rated higher in the intensity of sadness, happiness, and fearful emotions than male performers. On the other hand, male performers were assumed to be rated higher in the intensity of anger.

Additional Hypotheses According to Wester et al. (2002), Shields et al. (2006), and others, differences in ratings of both the male and the female subjects were assumed. Additionally, it was hypothesised that the ratings would differ even more for participants with a strong feminine or masculine side – according to the socially constructed gender, and not the biological gender.

Structure of the Paper

The thesis is divided into a part with general definitions followed by the three main sections and a concluding section. Each of the main sections are themselves divided in three sub-sections: (1) The method, (2) the results, and (3) the discussion. The first main section is a pre-experiment testing the videos and the type of questionnaire used on their applicability for the main study. Main section 2 moves on to describe the second pre-experiment, which was used for selecting the names. The second part is another pre-experiment, which is used to select names for the main study. Finally, section 3 discusses the main experiment: The main aims of the study were to link the fields of gender, emotion, and biological motion research, and explore the connection.

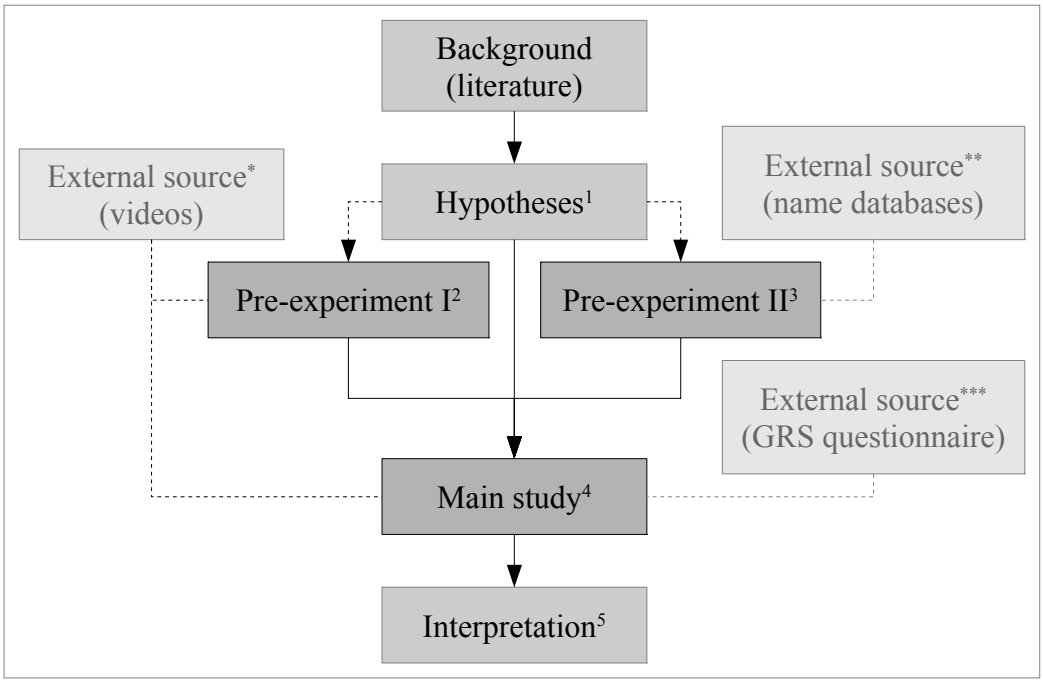


Figure 1. Structure of the paper.

Notes: [1]previous page (p. 6); [2]Pre-testing videos and questionnaire: method/results/discussion (pp. 9-24), appendices (pp. 91-122), digital appendices on CD (/content/experiments/pre-experiment_#1); [3]Name selection: method/results/discussion (pp. 25-36), appendices (pp. 123-189), digital appendices on CD (/content/experiments/pre-experiment_#2); [4]main study: method/result/discussion (pp. 37-75), appendices (pp. 190-235), digital appendices on CD (/content/experiments/main_study); [5]interpretation: discussion of the main experiment (pp. 65-75), and general discussion (pp. 76-84).

*provided by Mark Thompson, **online available (see additional references, p. 88), ***provided by Ursula Athenstaedt, PhD.

2 General Definitions

The most important definitions for understanding the paper are clarified in this section. More specific and detailed information can be found in the first paragraph of the sub-chapters describing the single steps (pre-experiments and main study) of the whole study.

Biological & Social Gender. In general[11], both terms – *biological gender* and *social gender* – are defined similar to Shields et al. (2006) throughout this paper: (1) *Biological gender* is the biological differentiation of women and men, and (2) *social gender* means the (culturally influenced) societal and/or behavioural aspects of the gender identity.

Female, Male, & Neutral gender. In addition to male and female *neutral gender* is used in the main study as well as the second pre-experiment. These terms describe names (not persons) and refer to prejudices and the cultural meaning – not to the biological gender.

Motion Capture Videos. This term describes the videos used as stimuli in this study. The *motion capture videos* show recordings of single points of pianist's upper body movement (the recordings are explained in the method section of the first pre-experiment, page 9)[12], whereby the points are connected through interpolated lines. The videos have no audio signal attached to them.

Motion. The term *motion* is used to describe changes in the position of the body, or just parts of it. *Motion* is used as synonym for *gestural expression* – not facial expressions.

Emotion. In this study, a discrete-category view[13] is used to distinguish *emotions*. Therefore *emotions* are qualitatively differentiated variants of affect (e.g. *anger*, *fear*, *happiness*, & *sadness*)[14], motivating more specific reactions in more specific situations.

11 Varying meanings are particularly marked.
12 Further information about the recordings can be found in Thompson and Luck (2008).
13 The models are shortly explained in the introduction (p. 2, for further reading see Thamm, 2006, Niedenthal, 2008).
14 Only the four exemplary mentioned emotions (anger, fear, happiness, & sadness) are used in this study.

3 Pre-Experiment I: Testing Applicability of Videos & Questionnaire

The main aim of the first pre-experiment was to determine if the videos are suitable for the main study: Primarily, the assumption that the videos are seen neutral (and *not* male or female) was tested. Two secondary objectives in addition to the main aim were set: (1) To determine if the text labels affects the reactions of the subjects, and (2) to test the interface of the digital questionnaire in handling and usability. Therefore, the subjects were presented 9 different motion capture videos showing 3 pianists performing the same piece of music in 3 different expression levels. All videos were labelled with the expression level intended by the performer.

Method

Stimulus. A total of 9 different videos were shown to the subjects. Marc Thompson recorded the videos for his own study[15] and kindly provided them for this study: Three pianists (2 females and 1 male) were asked to play the same piece of music[16] three times with different expressivity: Deadpan, normal, and exaggerated[17]. The terms used to describe the expression were chosen by Thompson and Luck (2008), and, therefore, also used in this study. The videos were taken using a point-light technique (PLT) to only record the movements without facial expression and body shape. In further preparation the videos were labelled on the bottom right corner with the expression intended therein (see Figure 2) using the open source software Kdenlive.

Material. The videos were recorded at the University of Jyväskylä (Finland). For recording the motion capture videos "[…] fifteen reflective markers were attached to key locations on the body (four on the head, one on each shoulder, one at the centre of the back, two on the lower back, two on the elbows, two on the wrists, one on each middle finger), and their three-dimensional spatial position recorded at 120 fps using an eight-camera optical motion capture system (Qualisys ProReflex). Additionally, two

15 The original videos were used for a research on expressivity of pianists (Thompson, & Luck, 2008).
16 In this case, not a whole piece of music was played: They played the first sixteen measures of the Brahms Intermezzo in A major Opus 118 #2.
17 Thompson and Luck (2008) explain that „the goal was to let each pianist interpret for him or herself what was meant by different levels of expression".

markers were placed at each end of the keyboard to act as reference points." (Thompson & Luck, 2008, p.541). The original motion capture videos were shortened to video clips showing the first 12 seconds only. Finally the videos were labelled and re-rendered using Kdenlive[18] on Ubuntu 10.04[19]. The render settings were: For video Xvid Version 4, 400k, 2 pass; without audio track added. The questionnaire for the experiment was digitalised and prepared for the realisation using PureData[20] on Ubuntu 10.04, and executed under Windows Vista[21] using Pure Data for Windows[22]. The videos were played using GEM[23].

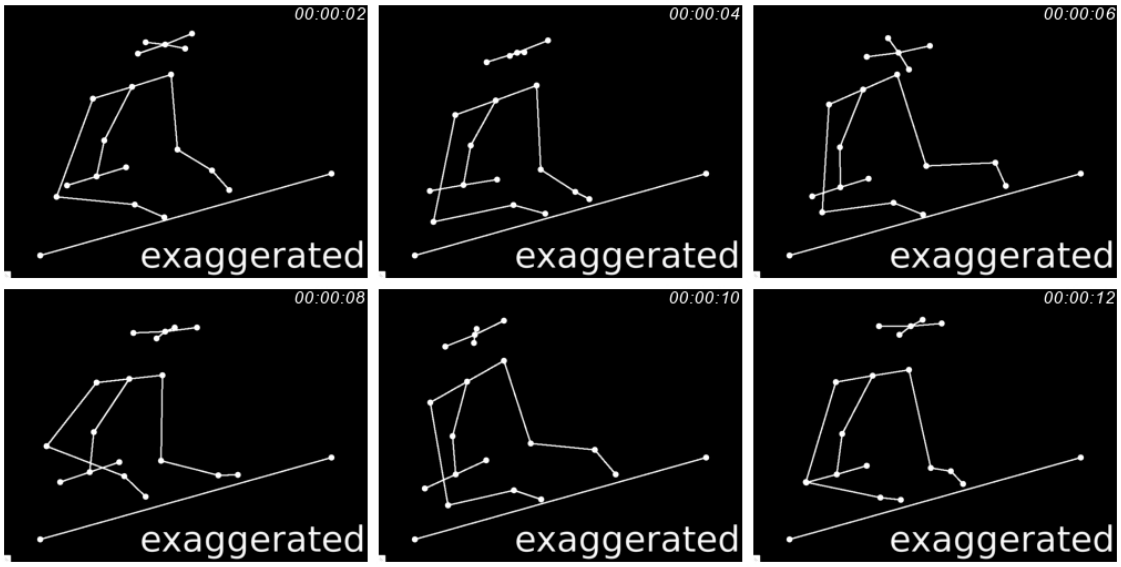


Figure 2. Exemplary frames of one motion capture video used in the study: The information provided to the participants was only visual – comparable to the image shown in the figure, only the movements and the expression were displayed (no further background information such as gender, or the piece of music which was played were given).

18 Used version: Kdenlive 0.9.2, using KDE Development Platform 4.4.5 [open source, online available: http://kdenlive.org]
19 Used version: Operating System – Ubuntu 10.04.4 LTS Lucid Lynx 32bit, Kernel 2.6.32-41-generic-pae [open source, online available: http://www.ubuntu.com]
20 Used version: Pd-extended 0.43.1 (*.deb) [open source, available on the appendix CD: /content/files/pd-extended_0.43.1/, online available: http://puredata.info/downloads]
21 Used version: Operating System – Windows Vista Service Pack 1, 32bit [Home Edition]
22 Used version: Pd-extended 0.43.1 (*.exe) [open source, available on the appendix CD: /content/files/pd-extended_0.43.1/, online available: http://puredata.info/downloads]
23 GEM stands for "Graphics Environment for Multimedia", used version: GEM 0.93.3 (implemented in Pd-extended) [open source, online available: http://gem.iem.at]

PRE-EXPERIMENT I: TESTING APPLICABILITY OF VIDEOS & QUESTIONNAIRE

Subjects. A total of 4 (3 male and 1 female) subjects volunteered to participate in the first pre-experiment. All were students and/or employees of the *Centre of Systematic Musicology* at the *University of Graz*. The subjects were between 23 and 54 years old (mean 34.75, standard deviation 13.99) with different subject-specific background in the field of investigation. Two subjects were professional music psychologists, the other two were students in the field of music psychology with profound knowledge of emotion research. The participants did not receive any incentive for their participation.

Procedure. The participants were briefed by the examiner that they will see computer animations of pianists while playing with different expression levels. After providing this information, the subjects were asked to rate the gender of the pianists shown in the video clips using a continuously variable horizontal slider from male (leftmost) to female (rightmost) as shown in Figure 3. In addition they had to rate (a) the duration of the whole study, (b) the comprehensibility of the instructions, and (c) to leave general comments.

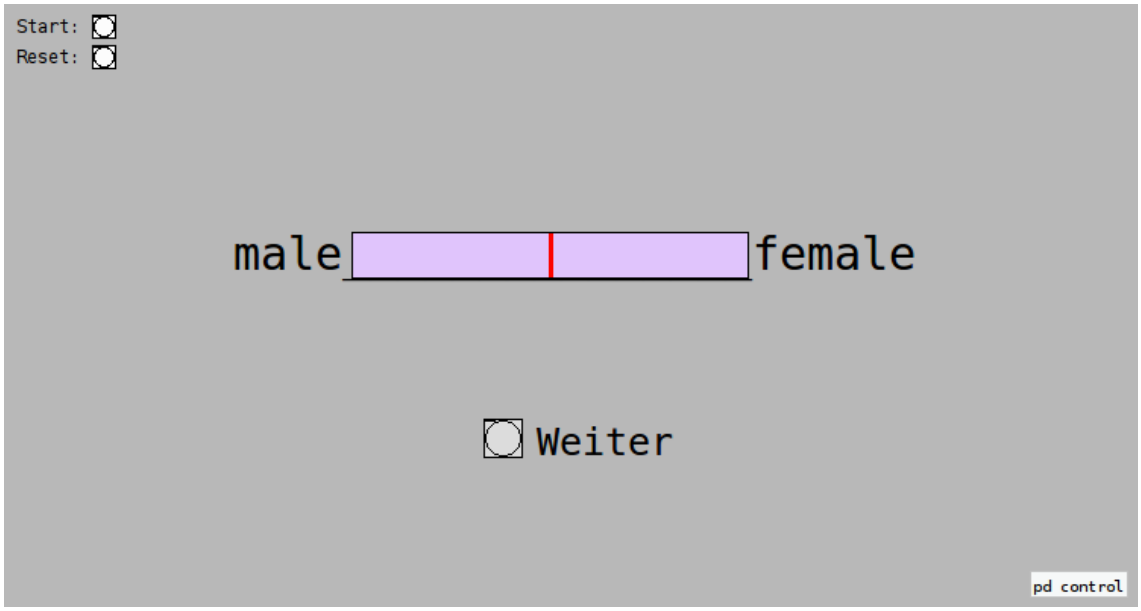


Figure 3. Screenshot of the input screen interface used in the first pre-experiment: The figure is showing the basic setting (slider on the value 0.0) presented to the participants right after viewing a video.

The main part of the experiment was conducted using a computer – the additional questionnaire and the demographic part were added in print-versions. For the computational part a patch[24] written in PureData[25] was used, which was on the one hand utilised to randomise the videos, and on the other hand used as a graphical input interface for the subjects (see Figure 3). Each of the 9 videos was shown to the participants 10 times. These 90 videos were then played in random order for each subject starting with one example[26]. The order was controlled by a pseudo-random generator built into PureData. The videos were embedded in the patch and played using GEM[27]. An example video started playing automatically after clicking the "Start"-button in the top left corner, and was shown in full screen mode. After each video, the participants were asked to rate the gender according their own perception (following the instructions given in the beginning of the experiment) using the input interface of the patch.

It was assumed that none of the videos can be assigned to a specific biological gender, or even identified[28] by the participants. In addition to this assumption, several reactions[29] on the text labelling were expected, whereby the term *reaction* wasn't defined precisely.

Results

Measure of Achievement. Primarily, the aim of the first pre-experiment was to test if the videos are suitable for the main experiment. In addition, two secondary objectives had to be verified: (1) To test if the text labels affect the ratings of the subjects, and (2) if the interface of the questionnaire is appropriate in handling and usability. Three different factors affected the measurement: First the quantitative data,

24 Technical details are described in appendix A, page 92**.**
25 PureData is an open source visual programming language. The used version can be found in the electronic appendix on the CD-ROM. Newer versions are available at http://puredata.info (official website).
26 The example was also randomly chosen out of the 9 videos but the input was not saved nor used for the analysis of data of the experiment. The example video was not part of the 90 videos shown to the participants.
27 GEM is an external used for PureData to implement graphics.
28 *Identifying* means assigning the real biological gender of the pianist playing in a video.
29 The interpretation of these reactions is explained in the next sub-section (Results).

second the (qualitative) comments/feedback by the subjects, and third the observation by the investigator. Due to the similarity in the knowledge and/or fields of research of the participants and the investigated field, the comments and the feedback play a greater role in the analysis of the data than the quantitative data.

However, the achievement of the major aim had to be measured by all three factors, but mainly by quantitative results, and the subjects' feedbacks. Therefore, a comparison of means was chosen for the analyses on the one hand, and on the other hand appropriate information of the participants' feedback was considered separately. Due to the small sample size, the analysis focused on the comments/feedback more than on the quantitative data. Indeed, this was only possible due to the background of the subjects volunteering in this pre-experiment – and their knowledge in (some of) the fields of investigation. Both secondary aims were adressed using the comments/feedback and the observation by the investigator only, without considering the quantitative data. Analysing if the text labels in the bottom right corner cause any reaction by the participants was primarily based on the observation by the investigator, and only slightly depended on the comments/feedback. On the contrary, the appropriate handling and usability of the input screen had to be analysed using the feedback and the comments by the participants.

The quantitative data was analysed in 3 steps: (1) A general analysis using an one-sample Student's *t*-test was conducted to compare the overall sample size mean ($\bar{\mathbf{x}}_{\mathbf{tot}}$) with the perfect neutral gender mean $\mu_0 = .00$, (2) the comparison of means of the two groups (students and non-students) was performed using Mann-Whitney U Test on the one hand, and an independent two-sample *t*-test on the other, and (3) the one-sample Student's *t*-test was performed for each of the two subsets' means ($\bar{\mathbf{x}}_{\mathbf{stud}}$, $\bar{\mathbf{x}}_{\mathbf{\neg stud}}$) separately.

Using a continuously variable slider on the computer resulted in float number values, which was good for analysing but not for displaying the data in bar charts and frequency tables. For that reason, all values were classified in steps of 0.2 as separate variables. The histograms and frequency tables are based on those classes, all analyses and the box plots are based on the original float number values.

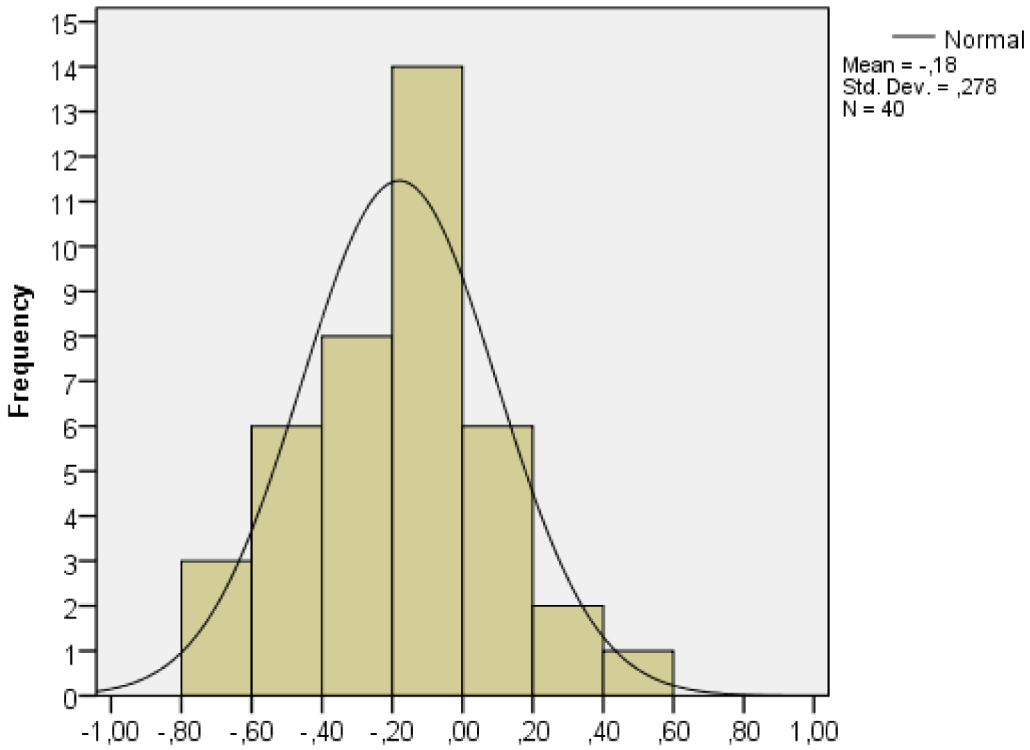


Figure 4. Histogram of gender ratings by the full sample (N = 40) for video no.7, classified in steps of 0.2.

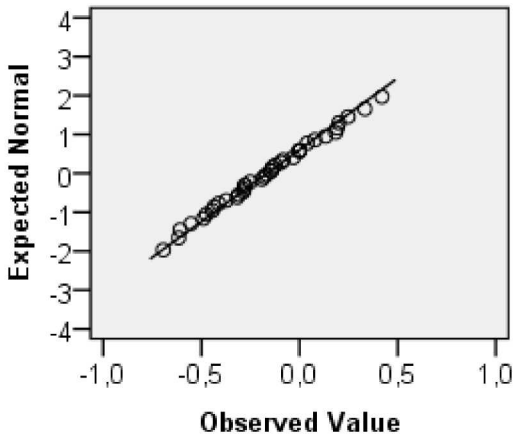


Figure 4a. Normal Q-Q Plot of gender ratings for video no.7.

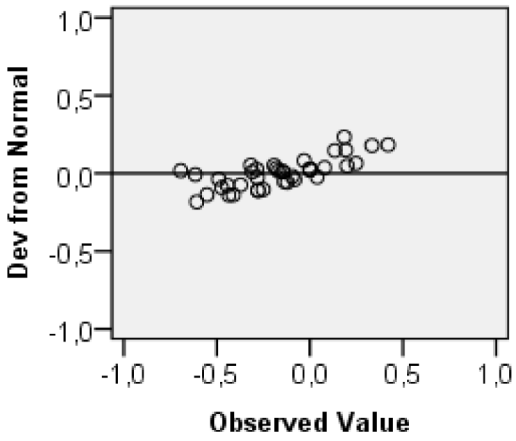


Figure 4b. Detrended Normal Q-Q Plot of gender ratings for video no.7.

Analysis of Data. The analysis of data will not go into detail regarding feedback and comments of the participants nor the observation by the investigator. This information is used in addition to the quantitative tests in the discussion section only.

Due to the computer based input, the complete quantitative data set was without any missing or invalid values. Therefore, the complete sample size of $N = 40$ (4 participants * 10 times per video) could be used for analysing. It was assumed that the data is normally distributed. To validate this assumption, a Kolmogorov-Smirnov Test and a Shapiro-Wilk Test were performed. The results showed that the data generally did not differ significantly from a normal distribution. One exception was found in the statistics of the Shapiro-Wilk Test; video no.4 ($W(40) = .924$, $p < .05$). The same result can be concluded from Figure 4.[30] In addition, the Normal Q-Q Plot in Figure 4a shows that the observed data follows a straight line; the same is true for the Detrended Normal Q-Q Plot (see Figure 4b).

The analysis of the general one-sample *t*-test showed several important results: As can be seen from Table 1, the mean ratings of the perceived gender are in a range between -.25 and .25. Those means were compared to $\mu_0 = .00$, which was defined (as mentioned before) to be the perfect neutral gender mean. A (two-tailed) one-sample *t*-test showed a highly statistically significant difference between the sample means and the hypothesised $\mu_0 = .00$ for video no.1 ($M = .233$, $SEM = .051$), $t(39) = 4.610$, $p < .001$; video no.5 ($M = -.184$, $SEM = .058$), $t(39) = -3.147$, $p < .01$; video no.7 ($M = -.164$, $SEM = .043$), $t(39) = -3.833$, $p < .001$; video no.8 ($M = -.215$, $SEM = .045$), $t(39) = -4.750$, $p < .001$; and a statistically significant difference for video no.2 ($M = .106$, $SEM = .048$), $t(39) = 2.199$, $p < .05$. The videos no.3, no.4, no.6, and no.9 did not statistically differ significantly from the perfect mean (see Table 2 for details). As can be seen from Table A.4-1 (Appendix A, page 103), approximately half of the videos showing a statistically significant difference have a higher value (videos no.1, & no.2; rated to be more female/feminine), and half to have a lower value (videos no.5, no.7, & no.8; rated to be more male/masculine) than the perfect neutral gender mean. Thus, a preliminary conclusion might be that it is possible that the gender ratings are totally random, but also could depend on the variable *student* (student versus non-student).

30 The figures for video no.7 are to be considered exemplary. All histograms and Q-Q Plots can be found in appendix A (Figures A-5.3 – A-5.11, page 114-122) as well as on the appendix CD.

Table 1

Means (M), Standard Deviations (SD), and Standard Error of Means (SEM) of the perceived gender for all test subjects.

	N	M	SD	SEM
video no.1	40	.233	.320	.051
video no.2	40	.106	.305	.048
video no.3	40	.077	.398	.063
video no.4	40	.006	.270	.043
video no.5	40	-.184	.369	.058
video no.6	40	.033	.326	.051
video no.7	40	-.164	.271	.043
video no.8	40	-.215	.286	.045
video no.9	40	.058	.412	.065

Note. N=40 (10 times each video per subject). The data is based on a continuous scale ranging from -1.0 (male/masculine) to 1.0 (female/feminine).

Table 2

Two-tailed one-sample t-test for $\bar{x}_{tot}$.

	Test Value = .00					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	M (MD)	Lower	Upper
video no.1	4.610***	39	.000	.233	.131	.336
video no.2	2.199*	39	.034	.106	.009	.203
video no.3	1.222	39	.229	.077	-.050	.204
video no.4	.149	39	.882	.006	-.080	.093
video no.5	-3.147**	39	.003	-.184	-.302	-.066
video no.6	.636	39	.528	.033	-.071	.137
video no.7	-3.833***	39	.000	-.164	-.251	-.078
video no.8	-4.750***	39	.000	-.215	-.306	-.123
video no.9	.896	39	.376	.058	-.073	.190

Note. N=40 (10 times each video per subject). The data is based on a continuous scale ranging from -1.0 (male/masculine) to 1.0 (female/feminine). Mean (M) and mean difference (MD) is the same for the test value $\mu_0 = .00$.

*p < .05, **p < .01, ***p < .001

The basic sample was divided in two groups to compare students' and non-students' ratings, both with a sample size of N = 20 (10 times each video was shown to 2 subjects). A Mann-Whitney *U* Test yielded that with an alpha level of .05 the two groups did not differ significantly in most cases; video no.1, $U = 186.5$, $Z = -.365$, p = n.s.; video no.2, $U = 130$, $Z = -1.894$, p = n.s.; video no.4, $U = 187.5$, $Z = -.338$, p = n.s.; video no.5, $U = 155.5$, $Z = -1.205$, p = n.s.; video no.6, $U = 166.5$, $Z = -.906$, p = n.s.; video no.8, $U = 187$, $Z = -.352$, p = n.s.; video no.9, $U = 160$, $Z = -1.082$, p = n.s.. A significant effect was found for video no.3, $U = 124.5$, $Z = -2.043$, $p < .05$, and video no.7, $U = 114.5$, $Z = -2.314$, $p < .05$.

Table 3

Mann-Whitney U Test comparing students and non-students.

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)	Exact Sig. [2*(1-tailed Sig.)]
video no.1	186.5	397	-.365	.715	.718[a]
video no.2	130.0	340	-1.894	.058	.060[a]
video no.3	124.5	335	-2.043	.041	.040[a]
video no.4	187.5	398	-.338	.735	.738[a]
video no.5	155.5	366	-1.205	.228	.231[a]
video no.6	166.5	377	-.906	.365	.369[a]
video no.7	114.5	325	-2.314	.021	.020[a]
video no.8	187.0	397	-.352	.725	.738[a]
video no.9	160.0	370	-1.082	.279	.289[a]

[a]Not corrected for ties.

In addition, an independent two-sample *t*-test was performed comparing the same two sub-samples. Levene's test for equality of variances showed no significant difference as can be seen from Table A-4.4 (page 107). As shown in Table 4, the *t*-test showed a statistically significant difference for video no.7 between the students' ratings (M = -.074) and the non-students' ratings (M = -.255), $t(38) = 2.206$, SED = .082, $p < .05$ (two-tailed). All other videos showed no significant difference between the two groups.

Indeed, this contrasts slightly with the outcome of the Mann-Whitney *U* Test, which showed a statistically significant difference for video no.3. However, no statistically significant differences between the two groups were found in general. Therefore, it can be concluded that the null hypothesis has to be accepted. No overall differences between students and non-students were shown by the quantitative data.

Table 4

Independent two-tailed t-test for equality of means of students and non-students.

	t	df	Sig. (2-tailed)	MD	SED	95% Confidence Interval of the Difference	
						Lower	Upper
video no.1	.710	38	.482	.072	.102	-.134	.279
video no.2	-1.608	38	.116	-.152	.094	-.343	.039
video no.3	-1.816	38	.077	-.222	.122	-.469	.025
video no.4	-.589	38	.559	-.051	.086	-.225	.124
video no.5	-1.045	38	.303	-.122	.117	-.358	.114
video no.6	-.380	38	.706	-.040	.104	-.250	.171
video no.7	2.206*	38	.034	.180	.082	.015	.346
video no.8	-1.235	38	.224	-.111	.090	-.293	.071
video no.9	1.317	38	.196	.170	.129	-.091	.432

*p < .05

Nevertheless, a (two-tailed) one-sample *t*-test was performed for both groups (students and non-students) separately. The sub-samples again were compared to the neutral value $\mu_0 = .00$. The tests yielded a statistically significant difference between the students' mean and the null value of $\mu_0 = .00$ with an alpha level of .05 in 3 cases (see Table A-4.7 on page 110 of the appendix for further information); video no.1 ($t(19) = 3.350$, SEM = .080, $p < .01$), video no.5 ($t(19) = -3.036$, SEM = .081, $p < .01$), video no.8 ($t(19) = -4.234$, SEM = .064, $p < .001$). At the same time, the null hypothesis could be rejected for the non-students' mean for 5 videos (see Table A-4.8 on page 111 of the appendix for further information); video no.1 ($t(19) = 3.154$, SEM = .062, $p < .01$), video no.2 ($t(19) = 2.435$, SEM = .075, $p < .05$), video no.3 ($t(19) = 2.914$, SEM = .064, $p < .01$), video no.7 ($t(19) = -5.076$, SEM = .050, $p < .001$), video no.8 ($t(19) = -2.521$, SEM = .063, $p < .05$). These results seem to stay in contrast to the outcome of the two-sample *t*-test.

Do the mean ratings of the videos not differ significantly from the perfect neutral mean (and tend to be neither seen as more female nor male)? The results for the whole sample supported this hypothesis in 4 out of 9 cases only (see Table 2). All other videos differed significantly from the neutral value. A similar result was found for the non-students group – the null hypothesis cannot be rejected for 4 videos out of 9. In contrast, the null hypothesis had to be rejected in 3 cases for the students group, only. Those findings partly support the assumption that the ratings do not differ significantly from the null value.

Do the means of the groups' ratings ("students" and "non-students") differ significantly from each other? Only minimal support was found for this hypothesis (see Table 3, and Table A-4.4). Statistical significance was achieved for video no.3 and video no.7 in the Mann-Whitney *U* Test and video no.7 in the *t*-test with an alpha level of .05. None of the other achieved statistical significance. However, different findings in the one-sample *t*-tests are strong points of supporting the hypothesis. Therefore, the data can't be clearly interpreted without analysing the feedback.

Discussion

This pre-experiment was designed to determine if the videos are suitable for the main experiment as well as to test the user interface in handling and usability. The analysis of the quantitative data showed that the mean ratings did not differ significantly from the perfect neutral mean, especially for the students' ratings. These findings validated the assumption that the pianists' gender cannot be identified by the participants.

It has to be questioned whether the difference between the mean ratings and the assumed null value is random or not. Based on the analysis of the quantitative data, it can be assumed that the findings are nonrandom, even if some of the data reached a significant difference from the neutral gender value. However, these results should be interpreted with caution. On the evidence presented, it cannot be certain whether the videos are seen as female, male, or neutral. This could be based on the differing views

on the terms. On the one hand, the students reported that "the points are all genderless to me" (own translation[31]), they only rated the gender "under compulsion" (non-literal translation/own translation[32]) caused by the slider, and questioned the gender concept used in the pre-experiment. On the other hand, the non-students commented that in context of the study they "do not know what can be interpreted as female [in this context]" (non-literal translation/own translation[33]) and see a direct connection between the amount of movement and the biological gender – "more movement [biological motion] = female?" (own translation[34]) – relating to the *lighter bodies* of women and/or *[male] homosexual stereotypes*. Figure 5 displays large variations between the individual ratings for most of the videos. Thus, due to the little impact on the gender rating, the videos can be seen as suitable for the main experiment.

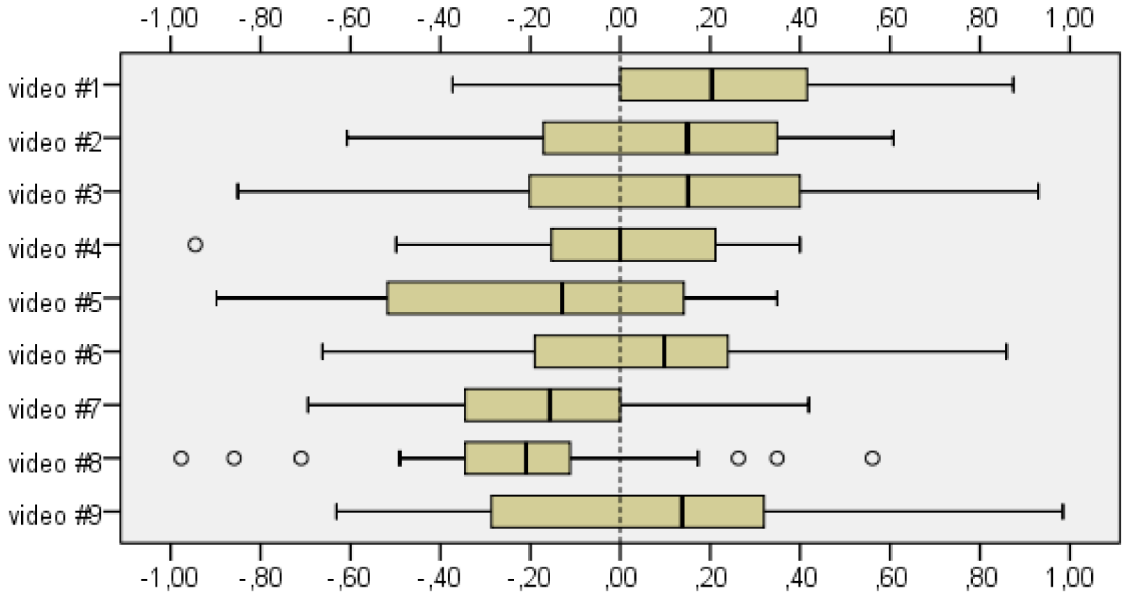


Figure 5. Box Plot for full sample (N = 40) gender ratings for all 9 videos.

Even if no statistically significant difference between the two groups (*students* and *non-students*) was achieved, the results for the one-sample *t*-tests revealed several differing results. As can be seen from Figure 6, the ratings differ slightly between the

31 Quoting the feedback originally written in German: "Die Punkte waren für mich alle geschlechtslos."
32 Quoting the feedback originally written in German: "Aus Zwang habe ich angekreuzt."
33 Quoting the feedback originally written in German: "Ich weiß nicht was 'weiblich' sein soll."
34 Quoting the feedback originally written in German: "Mehr Bewegung = weiblich?"

two groups: In average, the variations between the individual ratings are similar between both groups, but the students' means tend to be nearer to the neutral gender value than the non-students' ratings. In addition, the feedback showed that the students were questioning the gender in a different way (the recordings were genderless to them; without preferring one gender – neither male nor female) than the non-students (tending to rate the body weight as parameter of gender; based on stereotypes). Even if no statistically significant difference between the two sub-samples was found, a separate investigation for both was found to be more convenient in order to avoid bias.

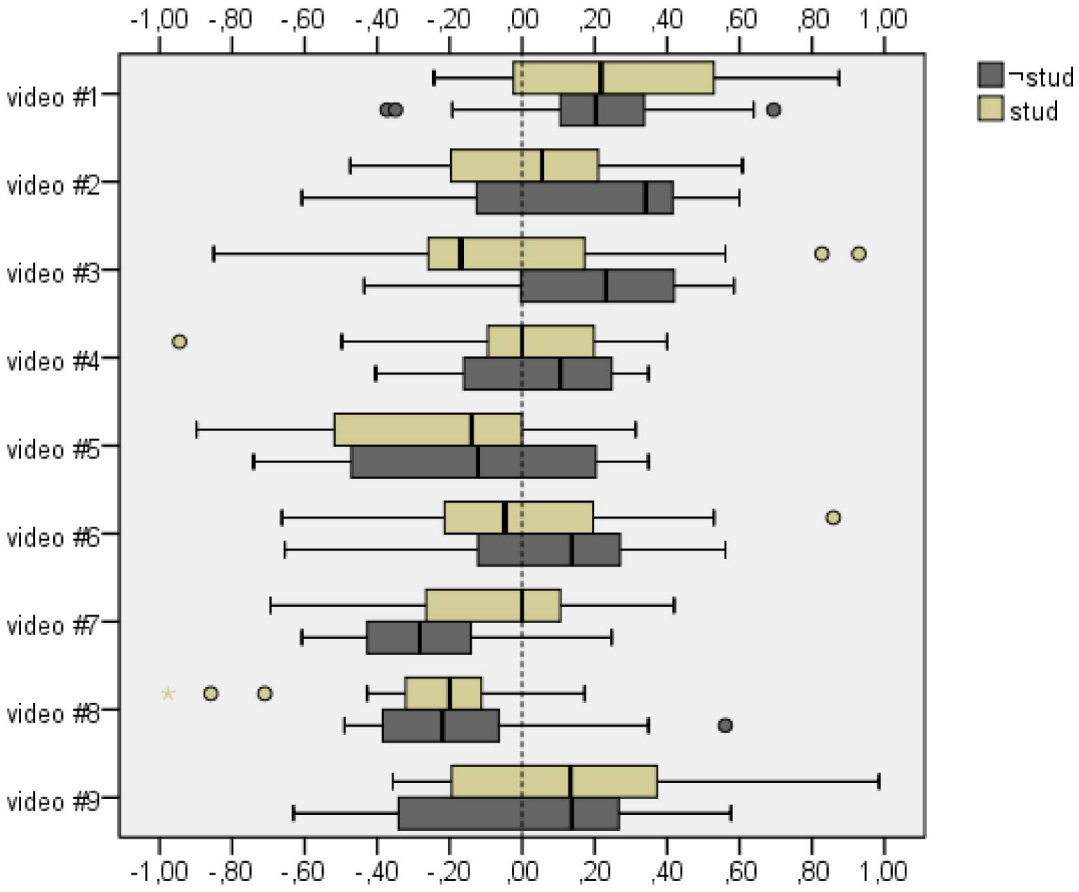


Figure 6. Box Plot comparing the two sub-samples, students (N = 20) and non-students (N = 20), with each other for all 9 videos.

Reactions obviously due to the text labels in the bottom right corner could be observed for all subjects. In all 4 cases, the participants reacted right after or already during the example video concerning the following statements[35]: If the label was put there by accident, and/or why the labels were put there, and/or what the meaning of the labels was. In addition, one feedback concerning the labels was written: "[D]eadpan, normal, exaggerated??" Therefore, no further observation by the investigator was required to decide whether the text labels affect the perception and the ratings of the subjects or not. Due to this obvious evidence, it can be assumed that the labels affect the perception and the ratings of the subjects.

The duration of the pre-experiment was rated as being too long in general by 3 of the 4 participants. "At the end [of the experiment] I couldn't figure any style of playing [any more], there were only points," (non-literal translation/own translation[36]) was only the most explicit example stated by one participant in the feedback sheet after finishing the experiment, another one wrote "good, but long" (non-literal translation/own translation[37]). However, the duration of the study lasted between 31 and 47 minutes in total, whereby the last 2 minutes were used for an informational dialogue and comments or statements on the pre-experiment. In conclusion, the video part of the main experiment had to be shortened.

Handling and usability of the used PureData patch were mainly measured by the physical reactions of the participants as well as their comments. One of the main problems in handling was the hardware – precisely, the handling of the mouse was affected by the surface of the table which caused problems in cursor movements. This could be observed in the movement of the hands moving the mouse. No other noteworthy problems in handling emerged. Several comments showed that the usability could be improved for example by adding a *repeat button*[38]. In addition, several comments on the study not important for further improvement were made: Several subjects criticised that no mean was displayed in the input screen to have a reference

35 One of the subject asked all three questions.
36 Quoting the feedback originally written in German: "Am Ende konnte ich keine Spielweise mehr ausmachen, es waren nur noch Punkte."
37 Quoting the feedback originally written in German: "[G]ut, etwas lang."
38 This was stated by two of the subjects during the experiment.

point for their ratings. Additionally most of them said that the possibility of resetting the slider to .00 would also improve the usability. One participant wrote that the "slider is not entirely clear" (own translation[39]) whilst another one wrote that the experiment was easily to understand in general. All in all, the results showed that a number of options (and partly the surface) had to be added to the PureData patch, adjusted or improved for a better usability.

One limitation of the first pre-experiment was the small sample size for each of the sub-groups. Therefore, the standard errors and confidence intervals can be larger than for large studies what makes it harder to interpret the p-values. Another limitation concerning the sample size is that, in fact, false positive results could be produced. On the other hand, all of the participants had a similar background to the field of investigation. Therefore, the comments and feedback could be used to support the results. Due to the small sample size the participants could be interviewed in greater detail than in large studies which can be seen as one of the major strength of a small sample size, especially useful in pre-experiments. A third limitation was caused by the reverse use of the *t*-test concerning the analysis: For results validating the hypothesis the null hypothesis couldn't be rejected. The problem was that the result only was based on the negotiation of the *t*-test.

To conclude, the videos were seen as neutral and therefore could be used as stimulus for the main experiment. The text label in the bottom right corner was noticed very well by the participants. Hence, no more adjustment of the font size and/or text position is required. Additionally, the results revealed a (not statistically significant) difference between students and non-students which was supported by the subjects' feedback. For that reason (to avoid bias), the main experiment was performed with students only. The pre-experiment also demonstrated that the viewing part of the study was too long (too many repetitions) and therefore was reduced for the main study. Furthermore, the suggested *repeat-button* was added for better usability.

39 Quoting the feedback originally written in German: "Schieberegler nicht ganz klar."

4 Pre-Experiment II: Selecting Names for the Main Study

For further preparation of the main study, the names used to label the videos had to be chosen: For this purpose a second pre-experiment was set up. The aim of this second part was to determine the associated (biological) gender of these names. This part of the study was set up without a hypothesis and/or scientific question, and was used as a decision support only. Nevertheless, the second pre-experiment was set up and analysed like a proper experiment.

Method

Stimulus & Material. For this pre-experiment 186 names[40] were preselected using several online name databases[41]. The chosen names were arbitrarily selected concerning their gender specific connotation in western culture. Although, names of different cultural origin were used, the originally gender-specific ascription[42] was not a decisive selection criterion for pre-selection. In a next step, the 186 names were listed in alphabetic order. This list was randomised[43] for each of the participants and set up as a questionnaire of 6 pages (31 names on each page). Unlike the other two parts of this master study, a hard-copy of the questionnaire on a sheet of paper was used to perform the second pre-experiment.

Subjects. A total of 10 (4 male and 6 female) subjects volunteered to participate in the second pre-experiment. All subjects participating in this pre-experiment had a similar cultural background[44]. The subjects were between 23 and 50 years old (mean 29, standard deviation 8.82) with different subject-specific background in the field of

40 Only first names were used in this study. The names are arranged alphabetically in appendix B, Table B-2.1 on page 125.

41 The used databases were: *http://www.beliebte-vornamen.de*, *http://www.namepedia.org*, *http://www.vornamen-weltweit.de*. *http://www.babynames.co.uk*, and *http://www.vornamen-online.ch* [all databases online available, date of access: March 5th 2012]

42 *Originally gender-specific ascription* refers to the gender ascription of the names based on the respective original cultural meaning.

43 The data was randomised using an online list randomiser [online available: http://www.random.org/lists/, date of access: April 17th 2012, and April 18th 2012]

44 The cultural background refers to the culture/country the participants grew up in.

investigation[45]. Seven of the subjects (3 male and 4 female) were students of the University of Graz and/or the University of Music and Performing Arts Graz. The participants did not receive any incentive for their participation. None of the subjects participated in the first pre-experiment.

Procedure. Before starting the experiment, the subjects were shortly briefed: They were instructed to rate the (biological) gender of names listed in the questionnaire according to their own opinion, starting with an example[46]. The gender of the names was rated on a five-step scale, as can be seen in Figure 7: *Female* (leftmost) – *mostly female – neutral gender – mostly male – male* (rightmost).

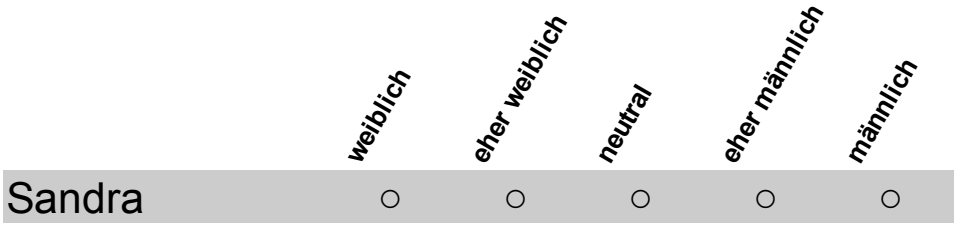


Figure 7. The figure shows the example on the first page of the questionnaire.

No specific outcome related to the gender of certain names was assumed, but it was expected that the findings show similarities due to the similar cultural background. Furthermore, it was neglected that slight differences with regard to other demographic parameters of the participants (such as biological or social gender, age, and education) could affect the findings.

45 The field of investigation refers on the one hand to the major research field of psychology, and on the other hand to onomatolgy, whereby none of the participants was familiar with the last mentioned.

46 The example name (Sandra) was not included in the analysis, nor was it included in the 186 names of the questionnaire.

Results

Measure of Achievement. Aim of this second pre-experiment was to find names that can be categorised very clearly concerning their (biological) gender – *female*, *male*, but also *gender neutral*. Therefore, a comparison of means was chosen for measuring the achievement, whereby this selection process was realised in 4 steps: (1) The mean of a name had to be in a range of .2[47] around the "perfect mean"[48]; (2) the highest standard error of mean allowed was predefined with 0.15[49]; (3) names that were ascribed to female, male or neutral (biological) gender in data bases were preferred; and (4) half of the chosen female, and half of the chosen male names were selected out of the best-liked names between 2010 and 2012[50].

Thus, the achievement of a name to be chosen for the main study was defined as a discrete function in dependence of the means ($\bar{\mathbf{x}}$) and the standard error ($\mathbf{y}$) of each name. Three achievement values were calculated for each name dependent on the perfect mean ($\mathbf{x_p}$). For female names x_p = -2.00, for neutral names x_p = ±.00, and for male names x_p = +2.00 was defined as perfect mean values. The achievement for each name $A(\bar{\mathbf{x}},\mathbf{y})$ was defined as

$$A_{i,j}(\bar{x},y) := \begin{cases} \dfrac{f_j(\bar{x})+f(y)}{2} & \text{if} \quad f(\bar{x}) \geqslant .95 \ \wedge \ f(y) \geqslant .95 \\ f_j(\bar{x}) \cdot f(y) & \text{if} \quad f(\bar{x}) < .95 \ \vee \ f(y) < .95 \end{cases}$$

$$f_j(\bar{x}) := \frac{1}{(\bar{x}_{i,j}-x_p)^2 \cdot C_1+1} \qquad f(y) := \frac{1}{y_i^2 \cdot C_2+1}$$

Two constant values were used to adjust the influence of means and standard errors on the total outcome: C_1 = 1.31, and C_2 = 2.33 (same for all cases). Due to the normalisation, the sub-functions – $f(\bar{\mathbf{x}})$, and $f(\mathbf{y})$ – may be any value between 0 and 1, whereby 1 was defined as the "ideal" value. Furthermore, for the achievement $A(\bar{\mathbf{x}},\mathbf{y})$ only positive values between 0 and 1 were possible. Equal achievement values are rated

47 For female names this range was set between -2 and -1.8, for neutral names between -0.2 and 0.2, and for male names between 1.8 and 2.

48 The "perfect means" were set for female -2, for neutral ±0, and for male +2.

49 Names with lower standard means were preferred.

50 The best-liked names were selected according to the used databases. All used databases are listed in the additional references (see page 88).

as equally strong, even if the means and standard errors differ from each other. Again, 1 was defined as the highest possible achievement, and all values for $A(\bar{\mathbf{x}},\mathbf{y}) \geq .95$ were accepted. A value below $A(\bar{\mathbf{x}},\mathbf{y}) = .95$ implies that either the mean deviates too far from the measured perfect mean, or the standard error is too high. In both cases, the null hypothesis had to be rejected.

In addition to the interpretation of means, the Student's *t*-test was used to validate the result. Therefore, the null hypothesis was tested for each of the three perfect means using the one-sample *t*-test. Each of the three tests was interpreted separately and compared with the analogous outcomes of the achievement function $A(\bar{\mathbf{x}},\mathbf{y})$ and the 4-step selection. Because of the formula of the Student's *t*-test no results with a standard deviation of .00 could be tested (division by zero).

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Analysis of Data. The data was assumed to be normally distributed. Due to the small sample size (N = 10), no additional test (like in the first pre-experiment) was performed to validate this assumption. No missing (or obviously invalid) values were found in the raw data. Therefore, the whole sample size was analysed for each name. The analysis section of the second pre-experiment is divided in 4 sub-sections: (1) An overall analysis of the data, (2) the analysis of data concerning the perfect female mean, (3) the analysis of data concerning the perfect male mean, and (4) the analysis of data concerning the perfect neutral mean.

Viewing the raw data in a first step revealed several interesting findings: Some of the names were only rated on the extreme values (left-most or right-most). As can be seen from Table 5, these names had a perfect mean with a standard error of SE = .00. The perfect matches were either male (M = 2.00) or female (M = -2.00) names, but none of them were neutral. For the perfect matching names no *t*-test could be performed because of the standard deviation of .00. Table B-3.1 (appendix B, page 125) shows that the data ranged between 0 (for the perfect matches) and 4.

Table 5

Perfect matching names in the categories male (M = 2.00, SE = .00, SD = .00) and female (M = -2.00, SE = .00, SD = .00).

	N	Range	M	SE	SD
Emilia	10	0	-2.00	.00	.00
Erik	10	0	2.00	.00	.00
Franz	10	0	2.00	.00	.00
Gerhard	10	0	2.00	.00	.00
Gustav	10	0	2.00	.00	.00
Hugo	10	0	2.00	.00	.00
Isabell	10	0	-2.00	.00	.00
Jana	10	0	-2.00	.00	.00
Johannes	10	0	2.00	.00	.00
Julia	10	0	-2.00	.00	.00
Lara	10	0	-2.00	.00	.00
Larissa	10	0	-2.00	.00	.00
Leyla	10	0	-2.00	.00	.00
Lukas	10	0	2.00	.00	.00
Markus	10	0	2.00	.00	.00
Matthias	10	0	2.00	.00	.00
Maximilian	10	0	2.00	.00	.00
Moritz	10	0	2.00	.00	.00
Nadine	10	0	-2.00	.00	.00
Nina	10	0	-2.00	.00	.00
Osman	10	0	2.00	.00	.00
Richard	10	0	2.00	.00	.00
Sven	10	0	2.00	.00	.00

Do the mean ratings not differ significantly from the perfect female mean ($\mu_0 = -2.00$); does the achievement value for $x_p = -2.00$ achieve the minimum required level ($A_{\bar{x},y} \geq .95$)? As can be seen from Table B-3.2 (appendix B, page 136), the null hypothesis did not have to be rejected for a total of 32 names ($A_{\bar{x},y} \geq .95$). The highest possible achievement level ($A_{\bar{x},y} = 1.000$) was measured for 9 names (Emilia, Isabell, Jana, Julia, Lara, Larissa, Leyla, Nadine, and Nina). Due to the standard deviation of SD = .00, no *t*-test could be performed to validate the result of these 9 names, nor was a validation necessary (all measured values were equal to the perfect mean). Furthermore, 15 names (Amelie, Bianca, Elena, Katharina, Lana, Lena, Leonie, Lucy, Mia, Natalie, Sarah, Sofia, Sophie, Susanne, and Vera) achieved the next highest value measured for $x_p = -2.00$ ($A_{\bar{x},y} = .982$, M = -1.90, range of 1). In addition to the achievement test, a (one-tailed) one-sample *t*-test was performed to reinforce these results. The test indicated that these 15 names ($t(9) = 1.000$, MD = .10, p = n.s., one-tailed)[51] do not differ significantly from the perfect female mean. Furthermore, the null hypothesis couldn't be rejected for 39 names in total (see Table B-3.2, appendix B page 136) - even for names not reaching the achievement level of $A_{\bar{x},y} \geq .95$.

The results supported the first hypothesis of the second pre-experiment (that any of the names are rated female) in 23 cases in both tests, achievement test and the one-tailed one-sample *t*-test with an alpha level of .05. Moreover, 9 names achieved the highest possible value in the achievement test ($A_{\bar{x},y} = 1.000$, M = -2.00, SE = .00). Due to the standard deviation of SD = .00, no additional *t*-test was performed. Therefore, the assumption that any of the names are rated female was validated for a total of 32 names. Additionally, the null hypothesis of the *t*-test could not be rejected for 16 further names (Alexandra, Alina, Anna, Astrid, Denise, Dina, Elise, Emma, Eva, Fiona, Florence, Mara, Maya, Rachel, Ronja, and Tanja). Even if these names were validated by the *t*-test, they did not reach the minimum required achievement level due to their wide dispersion.

51 Due to the same ratings for these names, the values are identical.

Do the mean ratings not differ significantly from the perfect male mean ($\mu_0 = +2.00$); does the achievement value for $x_p = 2.00$ achieve the minimum required level ($A_{\bar{x},y} \geq .95$)? As can be seen from Table B-3.3 (appendix B, page 144), the null hypothesis did not have to be rejected for a total of 35 names ($A_{\bar{x},y} \geq .95$). The highest possible achievement level ($A_{\bar{x},y} = 1.000$) was measured for 14 names (Erik, Franz, Gerhard, Gustav, Hugo, Johannes, Lukas, Markus, Matthias, Maximilian, Moritz, Osman, Richard, and Sven). Due to the standard deviation of SD = .00, no *t*-test could be performed to validate the result of these 14 names, nor was a validation necessary (all measured values were equal to the perfect mean). Furthermore, 16 names (Aaron, Adam, Arthur, Benjamin, Daniel, Felix, Jan, Justin, Klemens, Nikolaj, Oliver, Otto, Pedro, Simon, Tobias, and Tom) achieved the next highest value measured for $x_p = +2.00$ ($A_{\bar{x},y} = .982$, M = +1.90, range of 1). In addition to the achievement test, a (one-tailed) one-sample *t*-test was performed to reinforce these results. The test indicated that these 16 names ($t(9) = 1.000$, MD = .10, p = n.s., one-tailed)[52] do not differ significantly from the perfect male mean. Furthermore, the null hypothesis couldn't be rejected for 40 names in total (see Table B-3.3, appendix B page 144) - even for names not reaching the achievement level of $A_{\bar{x},y} \geq .95$.

The results supported the second hypothesis of the second pre-experiment (that any of the names are rated male) in 21 cases in both tests, achievement test and the one-tailed one-sample *t*-test with an alpha level of .05. Moreover, 14 names achieved the highest possible value in the achievement test ($A_{\bar{x},y} = 1.000$, M = -2.00, SE = .00). Due to the standard deviation of SD = .00, no additional *t*-test was performed. Therefore, the assumption that any of the names are rated male was validated for a total of 35 names. Additionally, the null hypothesis of the *t*-test could not be rejected for 19 further names (Andreas, Ben, David, Dennis, Elias, Ethan, Florian, Gideon, José, Julian, Leon, Matej, Mikkel, Nico, Patrick, Ryan, Sebastian, Tarek, and William). Even if these names were validated by the *t*-test, they did not reach the minimum required achievement level due to their wide dispersion.

52 Due to the same ratings for these names, the values are identical.

Do the mean ratings not differ significantly from the perfect neutral mean ($\mu_0 = \pm.00$); does the achievement value for $x_p = \pm.00$ achieve the minimum required level ($A_{\bar{x},y} \geq .95$)? As can be seen from Figure B-3.4 (appendix B, page 152), none of the names achieved the required value ($A_{\bar{x},y} \geq .95$). Therefore, the null hypothesis had to be rejected for all names for $x_p = \pm.00$. The *t*-test validated the outcome of the achievement test: The *t*-test yielded a statistically significant difference between $\mu_0 = \pm.00$ and the observed values mean with an alpha level of .05 for most of the names (127 out of 163 names[53]); 98 out of these 127 names even achieved a high statistically difference ($p < .001$). The mean ratings of the other 36 names did not achieve a statistically significant difference from the null value $\mu_0 = \pm.00$. Therefore, the null hypothesis couldn't be rejected for: Alani ($t(9) = -1.246$, MD = -.50, p = n.s.), Avery ($t(9) = .000$, MD = .00, p = n.s.), Ayse ($t(9) = -1.765$, MD = -.60, p = n.s.), Cameron ($t(9) = 1.500$, MD = .60, p = n.s.), Dace ($t(9) = 1.500$, MD = .40, p = n.s.), Diniz ($t(9) = -1.078$, MD = -.40, p = n.s.), Erin ($t(9) = -.519$, MD = -.30, p = n.s.), Ezra ($t(9) = -1.078$, MD = -.40, p = n.s.), Faizah ($t(9) = -2.090$, MD = -.70, p = n.s.), Gil ($t(9) = 1.861$, MD = .50, p = n.s.), Guilherme ($t(9) = .818$, MD = .30, p = n.s.), Hoa ($t(9) = 1.500$, MD = .40, p = n.s.), Irfan ($t(9) = 1.342$, MD = .50, p = n.s.), Jaime ($t(9) = 1.000$, MD = .40, p = n.s.), Janis ($t(9) = -.218$, MD = -.10, p = n.s.), Keiki ($t(9) = -.557$, MD = -.20, p = n.s.), Kim ($t(9) = -1.168$, MD = -.50, p = n.s.), Lærke ($t(9) = 1.309$, MD = .40, p = n.s.), Latif ($t(9) = 1.627$, MD = .50, p = n.s.), Madison, ($t(9) = -.768$, MD = -.40, p = n.s.), Makani ($t(9) = -.287$, MD = -.10, p = n.s.), Marit ($t(9) = -2.236$, MD = -1.00, p = n.s.), Maxime ($t(9) = -.885$, MD = -.40, p = n.s.), Mercedes ($t(9) = -1.203$, MD = -.60, p = n.s.), Nadim ($t(9) = 1.616$, MD = .60, p = n.s.), Pinar ($t(9) = -.218$, MD = -.10, p = n.s.), Pua ($t(9) = -.318$, MD = -.10, p = n.s.), Pualani ($t(9) = 1.000$, MD = .30, p = n.s.), Safa ($t(9) = -1.627$, MD = -.50, p = n.s.), Saga ($t(9) = -1.861$, MD = -.50, p = n.s.), Seher ($t(9) = 1.342$, MD = .50, p = n.s.), Sidney ($t(9) = -1.500$, MD = -.40, p = n.s.), Tuva ($t(9) = .000$, MD = .00, p = n.s.), Vega ($t(9) = .818$, MD = .30, p = n.s.), Vilde ($t(9) = .000$, MD = .00, p = n.s.), and Yael ($t(9) = 2.250$, MD = .60, p = n.s.).

53 In the experiment, 186 names were tested whereby 23 out of these names could not be acquired by the *t*-test due to the standard deviation of SD = .00.

The results of the third hypothesis that any of the names were rated neutral may be summarised as follows: The achievement values did not support the third assumption for any name. By contrast, the null hypothesis could not be rejected for 36 names in the (two-tailed) one-sample *t*-test. However, based on the comparatively high standard deviations (see Table B-3.4, appendix B page 152) for these names (minimum standard deviation for the sub-sample: $SD_{min} = .843$; maximum: $SD_{max} = 1.829$) this result must be interpreted with caution. Compared to the average standard deviation for the total sample ($\overline{SD}_{tot} = .667$) this sub-sample had a very high average standard deviation ($\overline{SD} = 1.144$), which shows that the ratings varied quite a bit.

Discussion

This pre-experiment was designed to find appropriate names to label the videos used in the main study. Therefore, the observed data was tested concerning three different hypotheses: Are the names ascribed to (1) female, (2) male, or (3) neutral gender. To sum up the findings of this second pre-experiment, two widely differing conclusions can be made: (1) Names that are well-known and marked by clichés and prejudices in the German-speaking areas can easily be associated with one specific biological gender (female or male), and (2) ratings for names that are rarely used and therefore not familiar in German-speaking areas scatter widely. These findings only partly support the name selection process.

Some of the results ought to be interpreted with caution (see Figure 8c) due to one strongly differing rating (this was the case for the names Anna, Eva, Patrick, Sebastian, Tanja, and William). At the same time, it can be, for instance, seen from Figure 8a (only female ratings) and Figure 8b (only male ratings) that at least some of the names were rated equal by all subjects.[54] However, no clearly discernible pattern was found in any of the questionnaires. Additionally, some of the names differed widely in their ratings as shown exemplary in Figure 8d.

54 Additional figures are in appendix B, Figure B-4.1 to Figure B-4.186, pages 160-183. All figures are also available on the appendix CD.

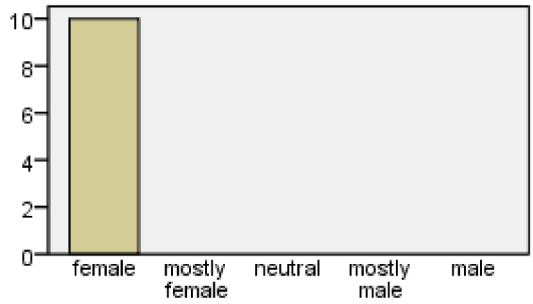


Figure 8a. Absolute number of ratings in each category for 'Nadine'

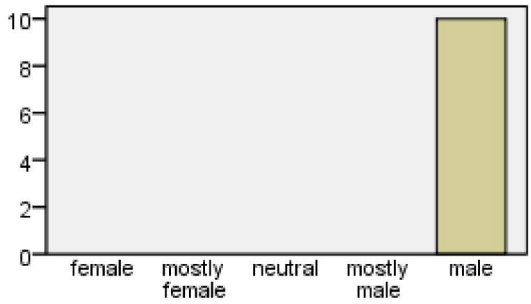


Figure 8b. Absolute number of ratings in each category for 'Johannes'

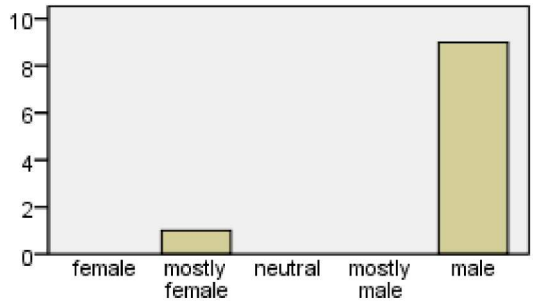


Figure 8c. Absolute number of ratings in each category for 'William'

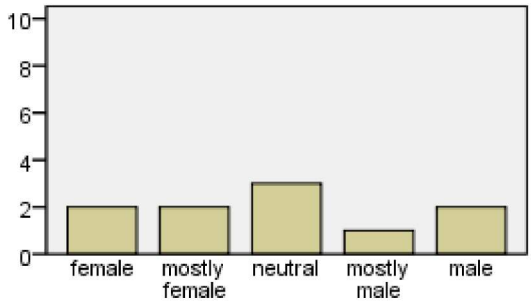


Figure 8d. Absolute number of ratings in each category for 'Janis'

The findings indicate that several female and male names are suitable to be used as stimuli in the main experiment. All in all, 32 names were associated with female gender and 35 with male gender. All of these names reached the required level of achievement, and in addition, none of these names differed statistically significant from the perfect means for female or male. Moreover, 9 of the names associated with female gender were observed to be considered as *perfectly female* (only rated as female); out of the 35 names associated with male gender, 14 were observed to be considered as *perfectly male* (only rated as male). In consideration of the first pre-experiment (due to the shortening of the video presentation part), only 18 names of both biological genders had to be selected for the main experiment. Therefore, 9 perfect matches and 9 *semi-perfect* matches (one rating mostly female/mostly male, all other ratings female/male) of both biological genders (female and male) were selected for the main experiment. All names were randomly selected out of the possible names.

By contrast, the hypothesis that any of the names were associated with none of both biological genders (neither female nor male) was only partly supported by the findings. On the one hand, none of the names was rated to be associated with *neutral gender* only, but on the other hand, no statistically significant difference from the perfect mean was observed concerning 36 cases. As can be exemplarily seen from Figure 8d, the results of the pre-experiment scattered widely for these names which results in a high standard error and a high standard deviation. Therefore, none of these names were selected as a stimulus for the main experiment. However, in addition to the main findings one additional interesting result was observed: All names partly validated to be associated with none of both biological genders had one thing in common: A different cultural background than the participants (non-German names, or *germanised* versions of non-German names). For all these reasons, the names used as neutral gender stimuli in the main study were not selected out of the preselected sample of names. Instead, the names were taken out of the category *neutral names* from the same databases used for pre-selection at first (the used databases are mentioned in the method section, page 25). In addition, the cultural background of the names was chosen to be the second selection criterion. The names were randomly selected out of the databases matching the given criteria.

Again, most of the limitations of the second pre-experiment were caused by the comparatively small sample size as well as in the first pre-experiment. Compared to the first pre-experiment, the limitations could not be compensated through additional comments, but the given sample showed that, usually, the opinions are culturally dependent. Another limitation of the second pre-experiment was the preselected sample of names used for testing. This limitation could be eliminated using the method of free association. Thirdly, the achievement test was strongly limited concerning neutral gender due to the strict rules. Although the additionally performed *t*-test was a little more open to differing values, the data scattered too widely to get clear results. On the one hand, the process of analysing could be improved by an alternative test for achievement, on the other hand, by a completely different experimental setup.

To sum up, several names were considered adequate for the main experiment. However, the names suitable for the main experiment were either female or male, whilst none of the pre-selected names was observed to be rated gender neutral. Hence, the female and male names used in the main experiment were chosen out of this pre-selected sample of names. For both, female and male names, 9 perfect matches and 9 additional (almost perfect) names were selected to be used in the main experiment. The following perfect matching names were selected: (a) For female: Emilia, Isabell, Jana, Julia, Lara, Larissa, Leyla, Nadine, and Nina; and (b) for male: Erik, Gerhard, Lukas, Markus, Matthias, Maximilian, Osman, Richard, and Sven. In addition, the following names were selected: (a) For female: Bianca, Elena, Katharina, Lana, Lena, Lucy, Mia, Natalie, and Sophie; and (b) for male: Aaron, Adam, Arthur, Benjamin, Daniel, Nikolaj, Oliver, Pedro, and Tom. The neutral names were chosen out of the same databases already used to preselect the names for this pre-experiment in accordance with the following two criteria: The names had to be (1) ascribed (in the database) to be a gender neutral name, or in other words, to be as likely female as male; and (2) non-German names, or *germanised* versions of non-German names (such as Sam, or Robin). On the basis of these criteria, the following neutral names were chosen out of the databases (randomly selected): Jamie, Lee, Maemi-Haru, Robin, Sam, Sasha, Summer, Uli, and Ying-Yu. The names chosen for the main study are alphabetically listed in Table B-3.1 (appendix B, page 133).

5 Main Experiment

The primary aim of this main part of the Master Thesis was to determine whether emotion is more strongly affected by biological gender or by motion. A number of researchers have already reported a very strong relation between the perception of emotions and the performer's biological gender (Hall, 1978, Behne, 1990, Locke, 2002, Brody & Hall, 2008). Therefore, it was hypothesised that the pianist's gemder has a stronger impact on the perception of emotion than expressive body movement does. Based on the assumption that the biological gender of a person is attributed through a name (due to prejudices), motion capture videos[55] (tested on applicability for the main experiment in the first pre-experiment) labelled with names (selected in the second pre-experiment) were presented to the subjects.

In addition, it was hypothesised that several other factors affect the ratings of the subjects: (1) The participants' gender was assumed to strongly affect their ratings; (2) the musical education was also considered to have an impact on the ratings of the subjects, whereby keyboard instruments (such as piano or keyboard) were assumed to have a greater impact than other instruments, and (3) several additional factors were assumed to play a minor role in influencing the outcome, like the gender or the social stratum of the participants[56]. The factors assumed to affect the ratings are supported by recent research (Locke, 2002, Wester et al., 2002, Fischer et al., 2004, Shields et al., 2006). However, the literature showed no tendency in any of the mentioned sub-theses. Therefore, the additional theses were investigated with a non-directional hypothesis.

55 The *motion capture videos* were played without audio signal. An exact explanation can be found in the definitions part of the paper on page 8.

56 This point was asked as last part of the demographic questionnaire, but answering it was not mandatory.

Method

Stimulus & Material. The main experiment was divided in two different major parts, using a stimuli in the first one, and a questionnaire in the second one: (1) The same 9 videos that were tested for applicability in the first pre-experiment (see page 9) were shown to the subjects, and (2) the participants had to answer a questionnaire concerning gender-role stereotypes. For the main study, the original videos were numbered consecutively as follows (the information about the real pianist's biological gender and the expression intended to be played was added enclosed in brackets)[57]: Video no.1 (female pianist no.1, deadpan expression), video no.2 (female pianist no.1, normal expression), video no.3 (female pianist no.1, exaggerated expression), video no.4 (female pianist no.2, deadpan expression), video no.5 (female pianist no.2, normal expression), video no.6 (female pianist no.2, exaggerated expression), video no.7 (male pianist, deadpan expression), video no.8 (male pianist, normal expression), video no.9 (male pianist, exaggerated expression). The videos were slightly adjusted for the main part: Instead of labelling the videos with the intended expression (as in the first pre-experiment), they were labelled with first names (selected in the second pre-experiment, see page 33) in the bottom right corner (one example is shown in Figure 9) using the same font, font size, and video editing software already used in the first pre-experiment. For the main experiment, the motion capture videos were shortened to video clips showing the first 15 seconds of the original videos. All videos were labelled and re-rendered using the same software already used for preparation of the first pre-experiment (see page 9). The names selected in the second pre-experiment were used to label the video clips. All names were assigned to the videos in the same order as listed[58]: The 9 gender neutral names were: Sam, Ying-Yu, Jamie, Robin, Maemi-Haru, Sasha, Uli, Lee, and Summer. The 18 female names were: Nadine, Lara, Jana, Isabell, Leyla, Larissa, Emilia, Julia, Nina, Mia, Sophie, Lucy, Lana, Katharina, Natalie, Elena, Lena, and Bianca. The 18 male names were: Gerhard, Richard, Markus, Osman, Maximilian, Matthias, Lukas, Erik, Sven, Arthur, Tom, Daniel, Aaron, Oliver, Pedro, Benjamin,

57 The numbering of the original videos was necessary for a clear assignment, because each label was only used once throughout the main experiment.

58 All names were randomised before assigning them to the videos. A list of how the names are combined with the labels can be found in appendix C, Table C-4.1, page 209.

Nikolaj, and Adam. A complete list of the videos (and how they were combined with labels) is added in appendix C (Table C-4.1, page 209). The same render settings were used for the videos, again without audio track added. In further preparation for easier handling in the PureData patch, the finished video clips were numbered consecutively starting with the 9 neutrally labelled video clips followed by the 18 video clips labelled with female names and the 18 video clips labelled with male names at last.

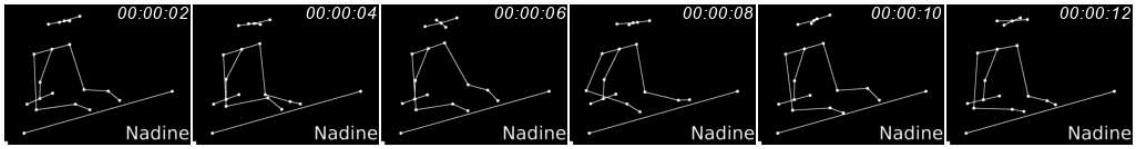


Figure 9. Video thumbnails of the first video clip labelled with a female name ('Nadine'). The recording was originally played by a female pianist in deadpan expression (original video no.1). The time code in the top right corner was additionally added in the thumbnails and not used during the experiment.

As discussed by Athenstaedt and Altstötter-Gleich (n.d.), research concerning gender attributes has relied mainly on self-descriptions (for further reading, see Athenstaedt, 2003, Saragovi et al., 2002). Therefore, a gender role attribute self-concept[59] was used to draw conclusions with regard to the subjects' gender for the second part of the main experiment. The "[s]ocially desirable [positive] and undesirable [negative] expressive [feminine] and instrumental [masculine] traits were measured with four scales" (Athenstaedt, 2003, p.312) in an adapted version of the German version of the Extended Personal Attributes Questionnaire (GE-PAQ, kindly provided by Ursula Athensteadt[60], University Professor at the Department of Psychology of the University of Graz; the German translated version based on Runge et al., 1981). In total, the GE-PAQ used in this study, consisted of 32 items, which could be divided evenly in four gender role attribute self-concept scales (each consisted of 8 items): The E+ (feminine/expressive and desirable/positive), the I+ (masculine/instrumental and

59 For detailed information about the gender role (attribute) self-concept see Athenstaedt (2003).

60 Additional information to the GE-PAQ was kindly provided together with the questionnaire in an unpublished manuscript (Athenstaedt, email, March 23, 2011). The questionnaire itself also was attached to the same email.

desirable/positive), the E- (feminine/expressive and undesirable/negative), and the I- (masculine/instrumental and undesirable/negative) scale. All items used for the positive (desirable) gender role attribute scales based on the GTS+ by Altstötter-Gleich (2004). "The items of the F+ [feminine/expressive and positive/desirable, E+] scale are: understanding, sensual, empathetic, romantic, soft-hearted, friendly/sincere, sensitive, and emotional[61]. The items of the M+ [masculine/instrumental and positive/desirable, I+] scale are: decisive, self-assured, fearless, assertive, self-confident, businesslike, daring, and commanding respect[62]" (Athenstaedt & Altstötter-Gleich, n.d., p.11). Furthermore, Athenstaedt and Altstötter-Gleich (n.d.) reported that they had to develop two new scales to measure the undesirable traits: "In the absence of a good instrument to measure the negative gender role attribute self-concept, we [Athenstaedt & Altstötter-Gleich] constructed two new scales […]." (p.10). The items of the F- [feminine/expressive and negative/undesirable, E-] are: whiny, weak, submissive, dependent, anxious, touchy, self-pitying, and always worrying[63]. The items of the M- [masculine/instrumental and negative/undesirable, I-] are: rude, competing, arrogant, dominant, egoistic, boastful, bossy, and consider oneself superior[64]. Figure C-1.2 in appendix C (page 192) shows a screenshot of the digitalised questionnaire used throughout the experiment. The Cronbach-α coefficients were calculated for all four scales as an indicator for their internal consistency. The coefficients were for the E+ scale $\alpha = .807$, for the I+ scale $\alpha = .875$, for the E- scale $\alpha = .832$, and for the I- scale $\alpha = .811$.

61 The original German terms used for the desirable expressive traits are (listed in the same order): verständnisvoll; sinnlich; einfühlsam; romantisch; weichherzig; herzlich; sensibel; gefühlsbetont.

62 The original German terms used for the desirable instrumental traits are (listed in the same order): entscheidungsfähig; trete bestimmt auf; unerschrocken; durchsetzungsfähig; selbstbewusst; zeige geschäftsmäßiges Verhalten, bereit, etwas zu riskieren; respekteinflößend.

63 The original German terms used for the undesirable expressive traits are (listed in the same order): weinerlich; schwach; unterwürfig; abhängig; ängstlich; empfindlich; selbst bemitleidend; bin ständig besorgt.

64 The original German terms used for the undesirable instrumental traits are (listed in the same order): rau; wetteifernd; überheblich; dominant; geltungsbedürftig; prahlerisch; diktatorisch; fühle mich überlegen.

In a third sub set, the demographic questionnaire was added followed by a questionnaire concerning the social strata of the participants. However, answering the last part of the experiment (the social strata questionnaire) was not mandatory. The questions for the questionnaire were taken from the appendix *'Dokumentation des Index der Sozialen Schicht'* (documentation of the social strata index) of the *16. Shell Jugendstudie* (edited by Shell Deutschland Holding, 2010, p.400f). According to the documentation of this study, the social strata is primarily based on the highest educational level of the family-father, and is differentiated using a self-assessment by adolescents concerning three aspects: (1) The financial situation, (2) the housing situation (house versus apartment, and owned property versus rented), and finally (3) the number of books (all three aspects referring to their parental household). Only participants with an age of 25 years and below were considered concerning the social strata; older students were not taken into account for the analysis.

Finally, all questionnaires were digitalised and embedded in the PureData patch. The main patch was structured in several sub-patches (for examples, see Figures 10-13). Each part of the experiment (input screen of the video presentation, gender role attribute self-concept questionnaire, demographic questionnaire, and social strata questionnaire) was put in an individual sub-patch and forced to close by the patch at the end of each separate part. During the main experiment, all sub-patches were shown to the participants on a 23 inch computer display in full screen mode. In contrast to the first pre-experiment, the PureData patch was executed under Windows XP[65] in the main experiment. Again, GEM was used to play the videos in the experiment. More detailed information about the PureData patch is added in appendix C1 (page 191).

Subjects. A total of 32 (12 male, 37.5%; and 20 female, 62,5%) subjects participated in the main part of this experiment. All participants were students of the University of Graz (68.7%), the University of Music and Performing Arts Graz (18.8%), or the Graz University of Technology (12.5%). The subjects were between 18 and 47 years old (mean 25.41, standard deviation 6.92) with different musical

65 Used version: Operating System – Windows XP Service Pack 3 [Professional Edition]

backgrounds. Most of the subjects were musically educated (a total of 29, 90.6%), and more than the half (a total of 17, 53.1%) mentioned to play a keyboard instrument (such as keyboard, piano, or similar). Only about two thirds of the participants was 25 years old or younger (a total of 21, 65.6%) and therefore included in the social strata statistics. Most of these 21 subjects (a total of 17, 81.0% of these 21 participants) stem from a high socio-economic background (8, 38.1% stem from the upper middle class, and 9, 42.9% stem from the upper class). The students were contacted via (1) e-mail, using several mailing lists of all three universities, (2) notices on university notice boards, and (3) a newsgroup posting. All participants received a small bag of chocolates as incentive, and had the chance to win a € 100,- gift coupon of a well known online retailer. None of the subjects participated in the first or second pre-experiment.

Procedure. The main experiment was conducted in two separate (similar) rooms by two investigators at the same time with overall duration of 20 – 30 minutes per participant (including the instructions). One of the investigators was female, the other one male in order to avoid bias (that the biological gender of the investigator has any impact on the participants' ratings). The experimental setup was the same, and the used hard- and software (screen size, input devices, both computers were Microsoft Windows based and the same PureData version was installed on both computers) was comparable in both rooms. With exception of a few instructions, the whole experiment was conducted using a PureData patch (with several sub-patches) on the computer.

Before executing the PureData patch on the computer to start the experiment, the subjects were orally instructed by one of the investigators and were informed that the experiment was divided in two major parts: (1) They will see short clips of motion capture recordings of piano players while playing, and (2) they will have to answer a short questionnaire. They were neither informed in detail what video clips were presented to them nor what the questionnaire was about. After this short introduction, the participants commenced the experiment by reading the instructions to the first part and clicking the *start button* (a screenshot of the introduction screen is shown in appendix C, Figure C-0.1, page 190) by themselves as soon as they were ready to start.

By starting the experiment, one of the example videos (see appendix C, Figures C-2.1 – C-2.8, page 198f) was randomly chosen, automatically opened in a GEM window, and presented to the subject. As soon as the video finished, the window was forced to close by the PureData patch and the input screen of the example was shown to the participant. The input screen contained the following elements (as shown in Figure 10): (1) The question on the top (repeated for each video slightly different by changing the mentioned name into the name used in the video clip): *'Which emotion(s) in which intensity would you assign to the piano playing of [name]?'*[66], (2) the rating scales for all four emotions: Sadness, fear, happiness, anger (same order as used in the study), (3) a *repeat button* (marked by the German phrase: 'Video erneut abspielen'), and (4) the *next button* (marked by the German phrase: 'Weiter zur nächsten Frage'). The intensities of all four emotions were rated on an 8-step scales from 0 (emotion is “non-existent” in the video) to 7 (“very strong/intense”). The ratings of the example were not considered in the analysis of the main experiment.

After finishing the introductory example, the participants were asked if there were any questions or obscurities about the experiment. In addition, the subjects were invited to test the usability and handling of the input screen (except the *next button*). After all questions and ambiguities were unambiguously settled, the subjects started the main part of the experiment by clicking the *next button*. The main part was set up identically to the introductory example: By clicking the *next button*, a new GEM window appeared and one randomly selected[67] video clip was presented to the participant. Each video was only played once (except for manually repetitions by the subject) and had to be rated according to the intensity of all four emotions afterwards.

66 The orignal German question is: 'Welche Emotion(en) würden Sie dem Klavierspiel von [Name] mit welcher Intensität zuordnen?'

67 The randomised selection was part of the PureData patch (for more detailed information, see appendix C1, page 191).

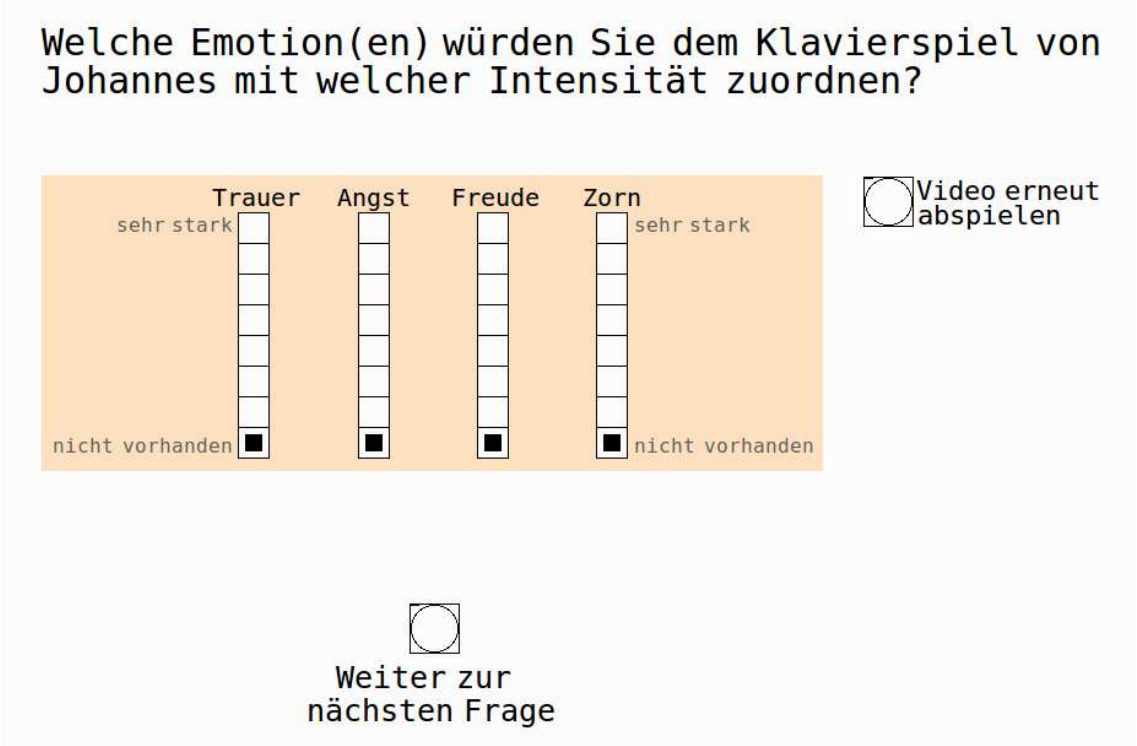


Figure 10. Input screen of the video presentation part in the main experiment (and the introductory example: The figure shows the basic setting presented to each subject after each video (all four emotions – f.l.t.r.: fear, happiness, sadnes, and anger - were preset to zero).

After all videos were presented to the participant once, the sub-patch showing the GE-PAQ (Figure 11) was started automatically. As explained above, the GE-PAQ is a system consisting of 4 scales and can be considered as one 4-dimensional, or two 2-dimensional systems (in this study, the latter structure was used[68]). As already mentioned in the Stimulus/Material section of the main experiment, each of the scales consisted of 8 items. All of the items were rated on a 6-step scale from 1 (“not correct at all”) to 6 (“applies completely”)[69]. Again, the subjects were shortly briefed by the investigator and all occurring questions were settled.

68 The 2 systems used in this study were E+/I+ (desirable traits) and E-/I- (undesirable traits).
69 Original German phrases for the scales for 1: “Trifft überhaupt nicht zu”, and 6: “Trifft völlig zu”.

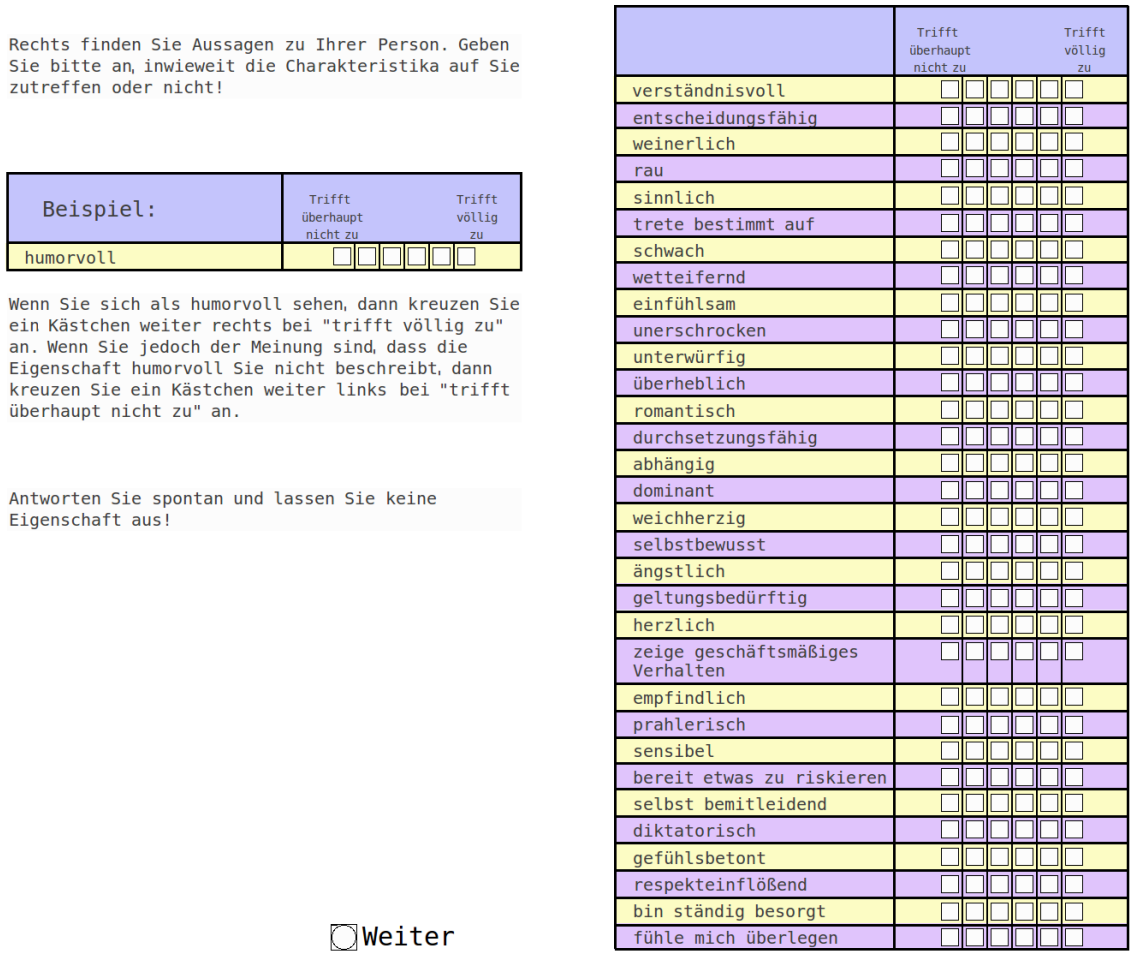


Figure 11. Screenshot of the adapted GE-PAQ used in the main experiment based on the printed version kindly provided by Dr. Ursula Athenstaedt (University of Graz, Austria).

An alert message pop-up window (see Figure 13) was used to avoid missing answers in the questionnaire. The message appeared if a participant clicked on the *next button* by accident without finishing the questionnaire (if no box was checked for one or more characteristics). In a third step, the sub-patch of the demographic data was shown to the participant (see Figure 12). The same alert message pop-up window already used in the GE-PAQ was used again to avoid missing answers for mandatory information.

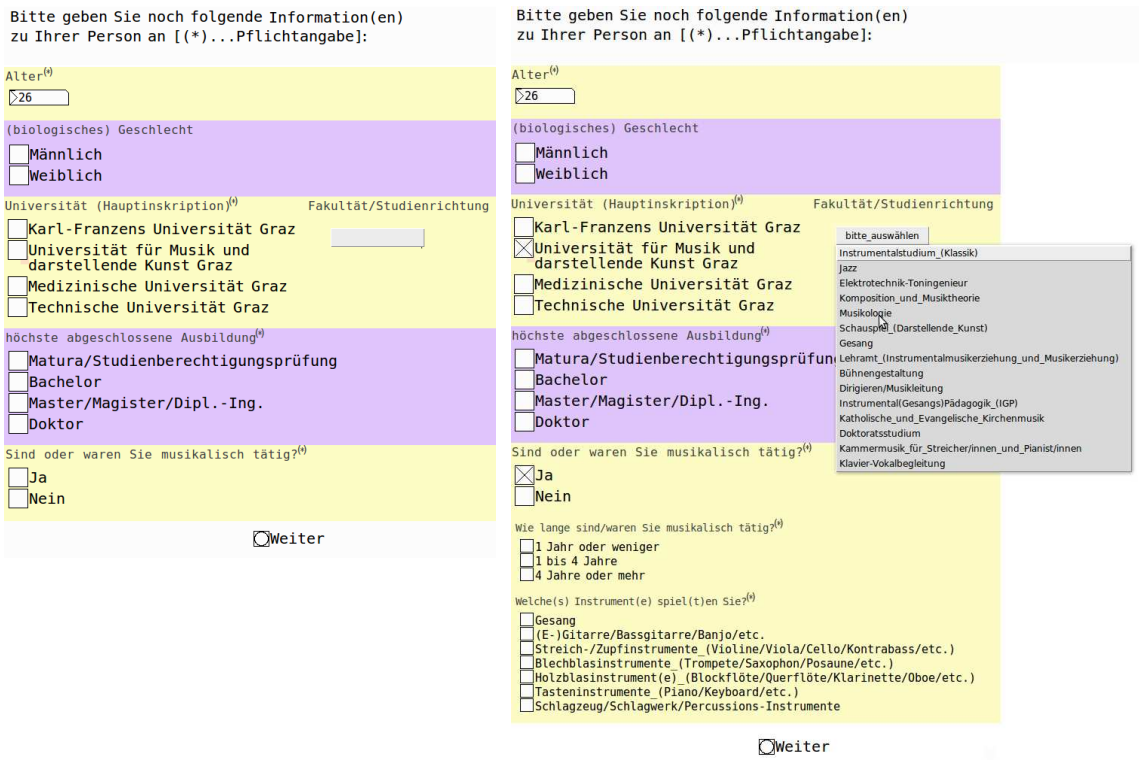


Figure 12. Screenshots of the demographic questionnaire PureData sub-patch: Basic setting (left), and additional options (right).

Finally, the sub-patch showing the social strata questionnaire was started. The participants were informed that answering it was not mandatory. Participants older than 25 years were told to skip the social strata questionnaire (because it is based on a study design prepared for adolescents from ages of 25 years and younger). The experiment was followed by an informal conversation about the whole study and the participants had the opportunity to ask questions in order to receive more information about the purpose of the experiment.

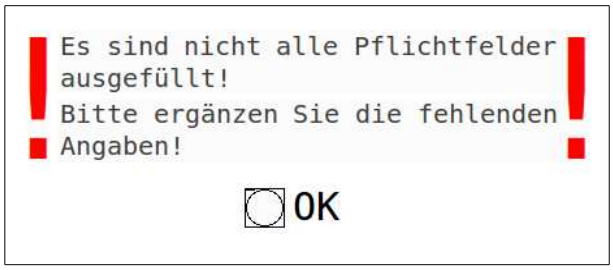


Figure 13. Alert Message Pop Up Window (PureData sub-patch) shown to the participants if any mandatory question was not answered.

Results

Measure of Achievement. The primarily aim of the main experiment was to investigate how biological gender and motion differ in their influence on the perception of gesturally expressed emotions. Both variables were assumed to affect how the intensity of emotions is perceived and therefore rated, whereby gender was assumed to affect the ratings more strongly than motion. In addition, several other factors were hypothesised to affect the results (undirected): The social gender (of the participant), the biological gender (of the participant), and the musical education (in particular whether a subject plays a key instrument or not). Furthermore, a number of additional possibly relevant factors were recorded: The real gender of the pianist shown in the recording (and if the true biological gender matches the stated gender), the age of the participants, the social stratum of the participant, and some additional information about the subjects' educational background (highest level of education, university, faculty and/or course of studies). The operationalisation was executed automatically as a part of the PureData patch and all additional values (for example if the real gender of the pianist matches the stated gender) were computed. The output was saved into a **.CSV* (comma separated value) text file and imported into an SPSS Statistics **.SAV* file.

In a first step, the items of the German Extended Personal Attributes Questionnaire (GE-PAQ) were combined to scales (as already noted in the methods section): Each scale was consisted of 8 traits (items) and always two scales formed a 2-dimensional system (one was positive or desirable, the second one negative or undesirable). In order to measure the achievement concerning the (social) gender, this simplification was necessary: The data was divided in the following four categories:

Feminine, masculine, androgynous, and undifferentiated. For example, as shown in Figure 14, a subject is androgynous if the following criteria are complied with: A high value on the expressive scale in combination with a high value on the instrumental scale; a subject is feminine if the value on the expressive scale is high and, at the same time, the value on the instrumental scale is low, and so on. In a second step, a combined system, including positive and negative trait scales, was computed. Therefore, each single value was weighted with the median:

$$x_E = \frac{x_{E+} \cdot \tilde{x}_{E+} + x_{E-} \cdot \tilde{x}_{E-}}{\tilde{x}_{E+} + \tilde{x}_{E-}} \qquad x_I = \frac{x_{I+} \cdot \tilde{x}_{I+} + x_{I-} \cdot \tilde{x}_{I-}}{\tilde{x}_{I+} + \tilde{x}_{I-}}$$

In this study 4 dependent variables (fear, anger, sadness, and happiness) were measured. As explained in the previous paragraph, each of the 2-dimensional systems of gender was reduced into one variable and accepted to be seen as independent variable for the analysis in the main section of this experiment. Therefore, the independent variables were: biological gender, intended expression, social gender (positive and negative), the musical education (in particular the ability to play the piano). A separate analysis of variances (ANOVA) was performed for each of the independent variables including the positive (desirable) and negative (undesirable) gender categorisations as well as the weighted overall gender categorisation. In addition, the directed hypotheses were tested performing a linear regression to analyse how well the independent variables explain changes in the dependent variables.

Even if the independent variables used in the linear regression were scaled categorically (expression: deadpan, normal, exaggerated; stated gender: female, male, neutral), they were accepted to fit for analysis. Therefore, all independent variables were encoded using a dummy-variable. The variable expression was encoded as follows: 0 = deadpan, 1 = normal, and 2 = exaggerated. The stated gender variable was encoded in a similar way: -1 = female, 0 = neutral, and 1 = male.

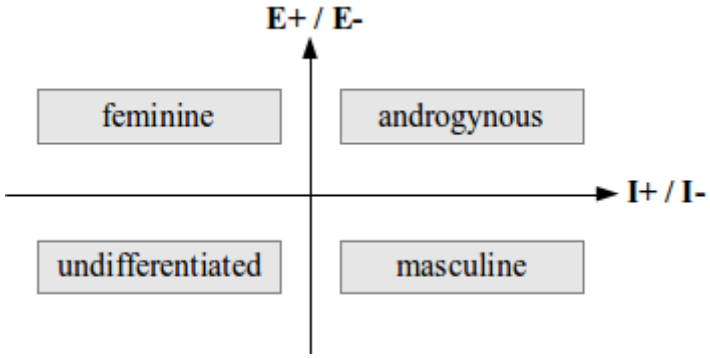


Figure 14. Interpretation of the 2-dimensional gender systems (combined graphic for positive and negative scaling).

Analysis of Data. The analysis of data is divided in 2 main parts: (1) The adjustment and operationalisation of the gender variables[70], and (2) the analysis of the main experiment. The second part is again structured into several sub-sections according to the main hypothesis and the sub-theses.

Due to the controlled (random) sample and test environment, no potential confounding variables were observed. However, the demographic data of the population (students) revealed tendencies according to gender and social strata of the subjects. These tendencies might be traced back to the fact that people with a higher level of education tend to be more androgynous than people with a lower level of education (information based on Ursula Athenstaedt, personal communication, March 30, 2011), and students usually stem from higher socio-economic backgrounds (Bundesministerium für Wissenschaft und Forschung, 2010, 2012). With this background knowledge, the gender variables were adjusted for the analysis:

First, the scales of the gender were computed. Therefore, the value for each scale was calculated determined by the mean of all 8 items. In a second step, the scales were combined into two 2-dimensional systems (positive and negative) as shown in Figure 14. As can be seen from Figure 15a and Figure 15b, the origin of the systems (3.5/3.5) had to be adjusted for categorisation[71]. The origin for the adjusted axes is equal

70 The only value not automatically computed by the PureData patch to operationalise the data for analysis was the gender categorisation.

71 Athenstaedt and Altstötter-Gleich (n.d.) used the same method for analysis of the German Extended Personal Attributes Questionnaire (GE-PAQ).

to the median of the scales: The median values were $\tilde{x}_{E+}=4.438$ and $\tilde{x}_{I+}=3.875$ for the positive scales, and $\tilde{x}_{E-}=2.500$ and $\tilde{x}_{I-}=2.750$ for the negative scales. Due to the overall tendency of university students to be androgynous, the categories were defined as follows: Higher or equal than the median for both (expressive and instrumental scale) was defined as *androgynous*; higher or equal than the median of the expressive scale, but lower than the median of the instrumental scale was defined as *feminine*; lower than the median on the expressive scale, and higher or equal to the median on the instrumental scale was defined as *masculine*; and finally, a value lower than the median on both scales was defined as undifferentiated:

$$\begin{aligned} \textit{androgynous} &:= x_E \geqslant \tilde{x}_E \ \wedge \ x_I \geqslant \tilde{x}_I \\ \textit{feminine} &:= x_E \geqslant \tilde{x}_E \ \wedge \ x_I < \tilde{x}_I \\ \textit{masculine} &:= x_E < \tilde{x}_E \ \wedge \ x_I \geqslant \tilde{x}_I \\ \textit{undifferentiated} &:= x_E < \tilde{x}_E \ \wedge \ x_I < \tilde{x}_I \end{aligned}$$

In a third step, the positive and negative scales were combined into one system. Therefore, the combined values were computed weighted with the median (as mentioned in the *measure of achievement* section). The Cronbach-α coefficients were calculated for both scales as an indicator for their internal consistency: The coefficients were for the (new) E_{tot} scale $\alpha = .689$, and $\alpha = .744$ for the (new) I_{tot} scale. The axes were adjusted the same way as for the other two gender scales: The median values were for the expressive scale $\tilde{x}_{E,tot}=3.717$ and $\tilde{x}_{I,tot}=3.443$ for the instrumental scale. The gender distribution of the overall system is shown in Figure 15.

The assumption that high values would be observed for instrumental as well as expressive items in a sample formed by students only was only partly supported. However, it was observed that desirable (positive) traits were rated high in most of the cases (see Figure 15a) whereby undesirable (negative) traits were rated comparatively low (see Figure 15b). In order to identify the trends concerning the gender, the data was re-coded into categories. Due to the unbalanced results, the axes were adjusted to get an approximately equal amount of subjects in each category.

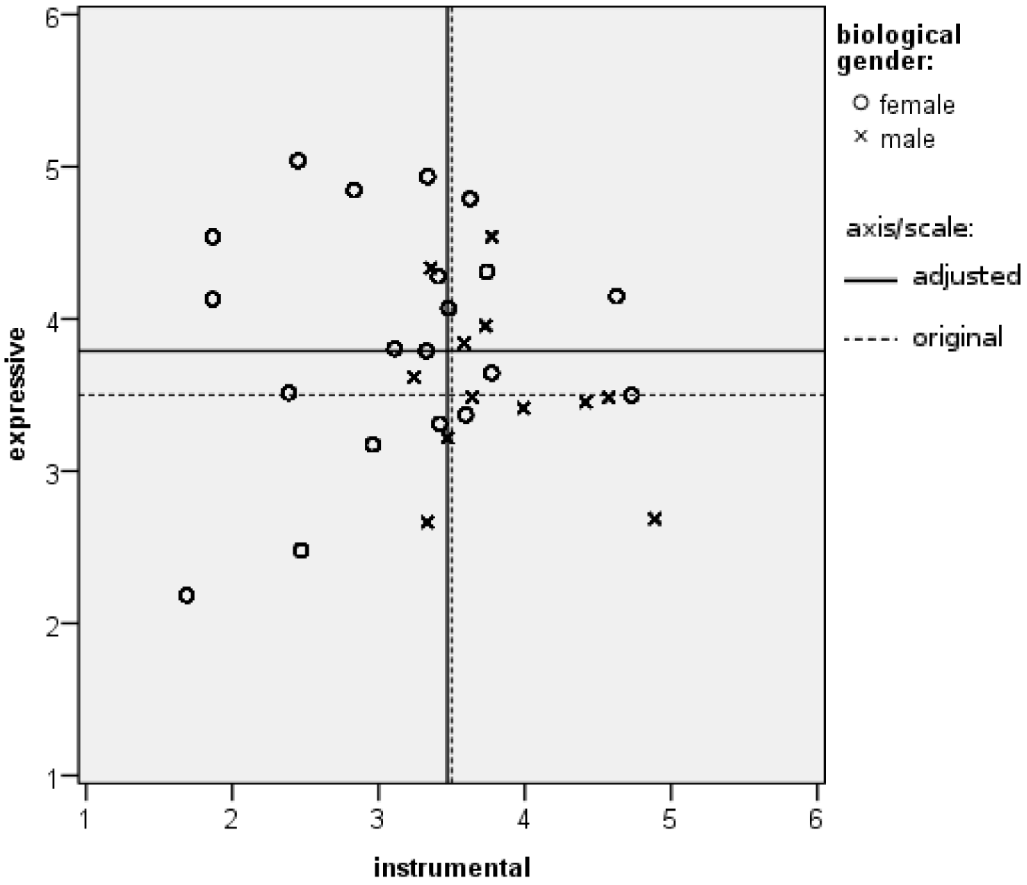


Figure 15. Gender distribution of the subjects depending on the weighted overall expressive and instrumental scale (weighted means).

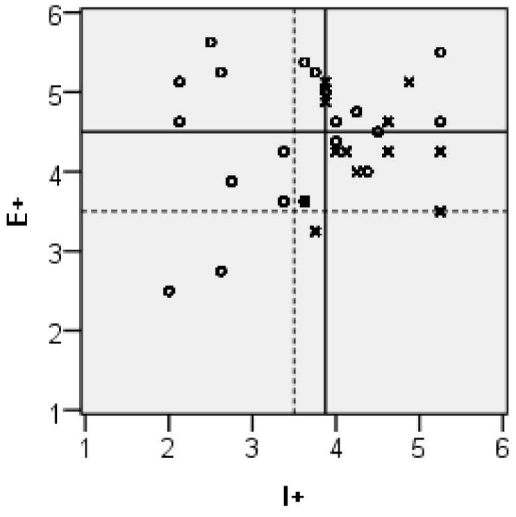


Figure 15a. Gender distribution depending on a positive expressive and instrumental scale.

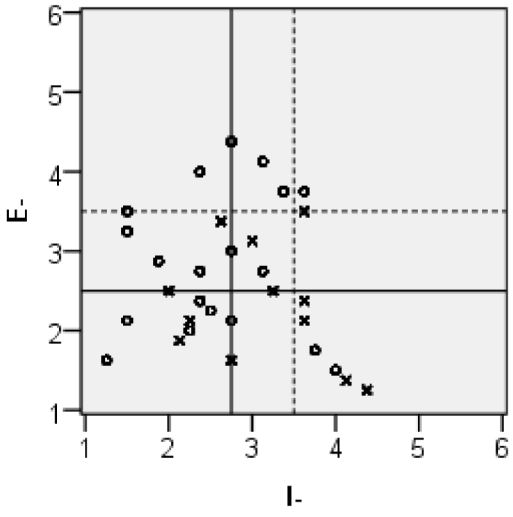


Figure 15b. Gender distribution depending on a negative expressive and instrumental scale.

To analyse the hypothesised influence of the pianist's (biological) gender and the intended expression, several tests were performed. In a first step, both independent variables were analysed conducting a one-way analysis of variance (ANOVA). Secondly, a linear regression was performed to analyse how strong a possible influence of the independent variables is:

First, an ANOVA was conducted to compare the effect of the stated gender on the intensity ratings concerning different emotions in female, male, and neutral conditions. The intensity of all four basic emotions (sadness, fear, happiness, and anger) observed in the experiment was considered separately. No statistical difference was found for any of the dependent variables: Fear, $F(2, 1437) = .255$, p = n.s.; happiness, $F(2, 1437) = .901$, p = n.s.; sadness, $F(2, 1437) = .742$, p = n.s.; and anger, $F(2, 1437) = .248$, p = n.s. (Table 6). For that reason, no post-hoc test was performed.

Table 6

Analysis of Variance (stated gender)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	1.626	2	.813	.255	.775
	Within Groups	4580.290	1437	3.187		
	Total	4581.916	1439			
happiness	Between Groups	7.590	2	3.795	.901	.407
	Within Groups	6054.899	1437	4.214		
	Total	6062.489	1439			
sadness	Between Groups	6.426	2	3.213	.742	.477
	Within Groups	6225.767	1437	4.332		
	Total	6232.194	1439			
anger	Between Groups	.800	2	.400	.248	.781
	Within Groups	2321.366	1437	1.615		
	Total	2322.166	1439			

Secondly, an ANOVA was conducted to compare the effect of the expression intended to be played by the pianist on the intensity ratings concerning different emotions in deadpan, normal, and exaggerated conditions. Again, the intensity of all four basic emotions (sadness, fear, happiness, and anger) observed in the experiment was considered separately. The intensity of emotions differed significantly across the three groups for each of the emotions: Fear, $F(2, 1437) = 109.666$, $p < .001$; happiness, $F(2, 1437) = 93.306$, $p < .001$; sadness, $F(2, 1437) = 8.578$, $p < .001$; and anger, $F(2, 1437) = 6.957$, $p < .01$ (Table 7).

Table 7

Analysis of Variance (intended expression)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	606.739	2	303.369	109.666	.000
	Within Groups	3975.177	1437	2.766		
	Total	4581.916	1439			
happiness	Between Groups	700.760	2	350.380	93.906	.000
	Within Groups	5361.729	1437	3.731		
	Total	6062.489	1439			
sadness	Between Groups	73.529	2	36.765	8.578	.000
	Within Groups	6158.665	1437	4.286		
	Total	6232.194	1439			
anger	Between Groups	22.268	2	11.134	6.957	.001
	Within Groups	2299.898	1437	1.600		
	Total	2322.166	1439			

Tukey post-hoc comparisons of the three groups were performed separately for each of the observed emotions: The comparisons for the emotion fear (see Table 8) indicate that the deadpan condition ($M = 2.14$, 95% confidence interval [1.95, 2.33]) was rated significantly higher than the normal condition ($M = 1.19$, 95% confidence interval [1.04, 1.33]), $p < .001$, and also significantly higher than the exaggerated condition ($M = .56$, 95% confidence interval [.46, .65]), $p < .001$. In addition, the comparisons suggest that the normal condition was rated significantly higher than the exaggerated condition, $p = .001$. For the emotion happiness (see Table 9), the Tukey post-hoc comparisons suggest that the deadpan condition ($M = 1.01$, 95% confidence interval [.87, 1.14]) was rated significantly lower than the normal condition ($M = 1.75$, 95% confidence interval [1.58, 1.92]), $p < .001$, and the exaggerated condition ($M = 2.71$, 95% confidence interval [2.51, 2.91]), $p < .001$. Furthermore, the comparisons indicate that the normal condition was rated significantly lower than the exaggerated condition. Comparisons for the emotion sadness (see Table 10) suggest that the deadpan condition ($M = 1.58$, 95% confidence interval [1.41, 1.76]) was rated significantly lower than normal condition ($M = 2.13$, 95% confidence interval [1.94, 2.33]), $p < .001$. Comparisons between the exaggerated condition ($M = 1.89$, 95% confidence interval [1.70, 2.08]) and the other two conditions were not statistically significant at a significance level of $p < .05$. For anger (see Table 11), the comparisons indicate that the exaggerated condition ($M = .79$, 95% confidence interval [.66, .92]) was rated significantly higher than the deadpan condition ($M = .57$, 95% confidence interval [.46, .68]), $p < .05$, as well as the normal condition ($M = .49$, 95% confidence interval [.39, .59]), $p < .01$. The comparison of normal and deadpan condition was not statistically significant at $p < .05$.

Table 8

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'fear'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			2.14	1.95	2.33
normal	.000*	1.000		1.19	1.04	1.33
exaggerated	.000*	.000*	1.000	.56	.46	.65
Total				1.29	1.20	1.38

*p < .001

Table 9

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'happiness'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			1.01	.87	1.14
normal	.000*	1.000		1.75	1.58	1.92
exaggerated	.000*	.000*	1.000	2.71	2.51	2.91
Total				1.82	1.72	1.93

*p < .001

Table 10

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'sadness'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			1.58	1.41	1.76
normal	.000*	1.000		2.13	1.94	2.33
exaggerated	.053	.167	1.000	1.89	1.70	2.08
Total				1.87	1.76	1.98

*p < .001

Table 11

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'anger'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			.57	.46	.68
normal	.613	1.000		.49	.39	.59
exaggerated	.022*	.001**	1.000	.79	.66	.92
Total				.62	.55	.68

*p < .05, **p < .01

In addition, a linear regression was performed to approximately evaluate how strong the influences of both independent variables on the perception of the intensity of emotions are. As shown in Table 12, the R-Squared values were low in general. However, the results of the regression indicated the two predictors (intended expression and stated gender) explained for the dependent variable fear 13.1% of the variance ($R^2 = .131$, $F(2,1437) = 108.039$, $p < .001$), for the emotion happiness 11.6% ($R^2 = .116$, $F(2,1437) = 94.252$, $p < .001$), for sadness 5% ($R^2 = .005$, $F(2,1437) = 3.319$, $p < .05$), and for anger also 5% ($R^2 = .005$, $F(2,1437) = 3.543$, $p < .05$). It was found that only the intended expression predicted the intensity of the perceived emotions with statistical significance, as can be seen from Table 13: Fear ($B = -.790$, $p < .001$), happiness ($B = .852$, $p < .001$), sadness ($B = .155$, $p < .05$), and anger ($B = .108$, $p < .01$). The stated gender did not reach statistical significance for any emotion at a significance level of $p = .05$. This outcome supports the findings of the ANOVA, even if the overall results predicted a low influence. However, in combination with the high statistical significance of the outcome concerning the independent variable 'intended expression', the results are highly convincing.

Table 12

R-Squared Table for the Model 'intended expression, stated gender'.

	R[a]	R Square	Adjusted R Square	SE of the Estimate
fear	.362	.131	.130	1.665
happiness	.341	.116	.115	1.931
sadness	.068	.005	.003	2.078
anger	.070	.005	.004	1.268

a. Predictors: (Constant), intended expression, stated gender.

Table 13

Linear regression analyses with intended expression and stated gender as independent variables for fear, happiness, sadness, and anger as dependent variables.

		Unstandardised Coefficients		Standardised Coefficients		
		B	SE	Beta	t	Sig.
fear	(Constant)	2.109	.099		21.346	.000***
	intended expression	-.790	.054	-.361	-14.695	.000***
	stated gender	-.023	.059	-.009	-.385	.700
happiness	(Constant)	1.074	.115		9.371	.000***
	intended expression	.852	.062	.339	13.671	.000***
	stated gender	-.087	.068	-.032	-1.273	.203
sadness	(Constant)	1.614	.123		13.091	.000***
	intended expression	.155	.067	.061	2.315	.021*
	stated gender	.083	.073	.030	1.132	.258
anger	(Constant)	.494	.075		6.562	.000***
	intended expression	.108	.041	.070	2.647	.008**
	stated gender	.013	.045	.007	.283	.777

*p < .05, **p < .01, ***p < .001

Besides the intended expression and the stated gender, all additional independent variables were analysed conducting a one-way ANOVA without performing additional tests. It was assumed that a strong influence would be observed for the participants' gender, their musical education (in particular keyboard instruments), and to a smaller extent the biological gender and the social stratum of each subject. In addition, the highest level of education was analysed. Due to the aforementioned homogeneity of level of education and social stratum among students, no significant differences were assumed. Therefore, these two aspects are only discussed briefly in this results section. For a better comparison between the social and the biological gender, the two analyses were first analysed separately, and compared in the following.

Therefore, an ANOVA was conducted to compare the effect of the participants' biological gender on the intensity ratings concerning different emotions of female and male participants. The intensity of emotions differed significantly across the three groups for the emotions fear, $F(1, 1438) = 14.471$, $p < .001$; happiness, $F(1, 1438) = 6.371$, $p < .05$; and anger, $F(1, 1438) = 47.080$, $p < .001$. No statistically significant difference was observed for sadness, $F(1, 1438) = .390$, $p = \text{n.s.}$, at a significance level of $p < .05$. No post-hoc test was performed because there were fewer than three groups/conditions (female and male). However, the test indicates that male participants tended to rate emotions higher in intensity in general: Even if the emotion anger was rated low in general by males (M = .91, 95% confidence interval [.78, 1.03]) as well as females (M = .44, 95% confidence interval [.37, .51]), the relative difference was observed to be the highest of all emotions. The second highest relative difference was observed for the emotion fear: Male participants (M = 1.52, 95% confidence interval [1.37, 1.67]) and female participants (M = 1.15, 95% confidence interval [1.04, 1.27]). Happiness did not differ as strongly between the males' (M = 2.00, 95% confidence interval [1.83, 2.17]) and the females' (M = 1.72, 95% confidence interval [1.58, 1.85]) ratings as anger and fear. Moreover, even if sadness did not reach statistical significance, it was shown that male participants (M = 1.91, 95% confidence interval [1.75, 2.08]) rated slightly higher in intensity than female participants (M = 1.84, 95% confidence interval [1.70, 1.98]).

Furthermore, an ANOVA to compare the effect of the participants' gender on the intensity ratings concerning different emotions in androgynous, feminine, masculine, and undifferentiated conditions was conducted once for gender categories of desirable traits, once for undesirable traits, and once for the weighted overall gender categories. The intensity of all four basic emotions (sadness, fear, happiness, and anger) observed in the experiment was considered separately with different outcomes for the social gender categories according to desirable and undesirable personality traits, as well as the overall weighted gender categories:

Social gender categories according to desirable (positive) personality traits: Statistically significant differences across the 4 conditions (feminine, masculine, androgynous, and undifferentiated) were found for fear, $F(3, 1436) = 3.436$, $p < .05$; and anger, $F(3, 1436) = 5.499$, $p < .01$; whilst no statistically difference was found for the dependent variables happiness, $F(3, 1436) = .995$, p = n.s.; and sadness, $F(3, 1436) = 2.336$, p = n.s.

Social gender categories according to undesirable (negative) personality traits: The results of the social gender categories according to desirable personality traits were not consistent with the findings of those according to undesirable personality traits. Again, statistically significant differences across the 4 conditions (feminine, masculine, androgynous, and undifferentiated) were found, but under these conditions for fear, $F(3, 1436) = 4.512$, $p < .01$; and sadness, $F(3, 1436) = 4.539$, $p < .01$. No statistically difference was found for the dependent variables happiness, $F(3, 1436) = 1.879$, p = n.s.; and anger, $F(3, 1436) = .586$, p = n.s.

Weighted overall social gender categories: Three of the four emotions differed with statistical significance across the four conditions feminine, masculine, androgynous, and undifferentiated: Fear, $F(3, 1436) = 17.620$, $p < .001$; sadness, $F(3, 1436) = 18.113$, $p < .001$; and anger $F(3, 1436) = 10.499$, $p < .001$. Happiness, $F(3, 1436) = .684$, p = n.s., did not reach statistical significance. Tukey post-hoc comparisons were performed for fear, sadness, and anger.

Table 14

Tukey post-hoc comparisons of the weighted overall (social) gender categories for the dependent variable 'fear'.

						95% Confidence Interval for Mean	
	fem.	masc.	andr.	undif.	M	Lower Bound	Upper Bound
feminine	1.000				.86	.73	.99
masculine	.000**	1.000			1.54	1.36	1.72
androgynous	.000**	.641	1.000		1.70	1.45	1.95
undifferentiated	.195	.007*	.000**	1.000	1.12	.95	1.29
Total					1.29	1.20	1.38

*p < .01, **p < .001

Table 15

Tukey post-hoc comparisons of the weighted overall (social) gender categories for the dependent variable 'sadness'.

						95% Confidence Interval for Mean	
	fem.	masc.	andr.	undif.	M	Lower Bound	Upper Bound
feminine	1.000				1.50	1.32	1.68
masculine	.494	1.000			1.70	1.53	1.88
androgynous	.000*	.000*	1.000		2.58	2.31	2.86
undifferentiated	.126	.818	.000*	1.000	1.84	1.61	2.06
Total					1.87	1.76	1.98

*p < .001

Table 16

Tukey post-hoc comparisons of the weighted overall (social) gender categories for the dependent variable 'anger'.

						95% Confidence Interval for Mean	
	fem.	masc.	andr.	undif.	M	Lower Bound	Upper Bound
feminine	1.000				.47	.36	.58
masculine	.006*	1.000			.76	.63	.88
androgynous	.000**	.760	1.000		.85	.67	.103
undifferentiated	.855	.001*	.000**	1.000	.39	.29	.50
Total					.62	.55	.68

*p < .01, **p < .001

The comparisons indicate that androgynous individuals rated the intensity of all three emotions (fear, sadness, and anger) significantly higher than individual with feminine or undifferentiated gender at a significance level of $p < .05$, as can be seen from Table 14, Table 15, and Table 16. Masculine individuals did not differ significantly from androgynous individuals in rating the intensity of fear and anger, but they rated the intensity of sadness significantly lower than androgynous individuals: Masculine ($M = 1.70$, 95% confidence interval [1.53, 1.88]), androgynous ($M = 2.58$, 95% confidence interval [2.31, 2.86]), $p = .000$. The comparison of undifferentiated and feminine individuals was not significant at $p < .05$ for any of the emotions.

Another ANOVA was used to test for differences between individuals with musical background and without musical background. A statistically significant difference between musicians and non-musicians was found for the emotion sadness, $F(1, 1438) = 8.847$, $p < .01$, but not for any of the other three emotions (not significant for fear, happiness, and anger). As expected, no statistical significance was reached testing the differences among individuals who could play a keyboard instrument, and those who couldn't.

In addition to the main assumption that the stated gender affects the perception of intensity ratings about emotions, the real biological genders of the pianists performing in the videos were tested for differences with regard to how the emotional content is perceived (rated). Even if the real biological gender was not visible to the participants, the ANOVA indicated that it affected the ratings of the participants. Intensity ratings differed significantly across both biological genders (see Table 17) for all for dependent variables: Fear, $F(1, 1438) = 5.940$, $p < .05$; happiness, $F(1, 1438) = 36.225$, $p < .001$; sadness, $F(1, 1438) = 74.791$, $p < .001$; and anger $F(1, 1438) = 95.071$, $p < .001$. Videos performed by a male pianist ($M = 1.21$, 95% confidence interval [1.06, 1.37]) were rated significantly lower in intensity than videos performed by a female pianist ($M = 2.20$, 95% confidence interval [2.06, 2.33]), $p < .001$, for sadness only. All other emotions were rated significantly higher in intensity for videos performed by a male pianist. The highest relative difference was (again) observed for the emotion anger: Male pianist, $M = 1.06$, 95% confidence interval [.92, 1.21]; female pianist, $M = .39$, 95% confidence interval [.33, .46]; $p < .001$. However, this finding supports the assumption, that the biological gender affects the perception of emotion.

Due to this finding, another ANOVA was performed concerning the combination of the real gender of the pianists and the stated gender: The test was conducted to ascertain if the result was also significantly different for videos where the stated gender matched the real gender and those where both biological genders didn't match. However, no statistically significant difference was found among the different conditions for any of the emotions (fear, happiness, sadness, and anger).

Table 17

Analysis of Variance (real biological gender of the pianist)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	18.850	1	18.850	5.940	.015
	Within Groups	4563.066	1438	3.173		
	Total	4581.916	1439			
happiness	Between Groups	148.967	1	148.967	36.225	.000
	Within Groups	5913.522	1438	4.112		
	Total	6062.489	1439			
sadness	Between Groups	308.113	1	308.113	74.791	.000
	Within Groups	5924.081	1438	4.120		
	Total	6232.194	1439			
anger	Between Groups	144.006	1	144.006	95.071	.000
	Within Groups	2178.160	1438	1.515		
	Total	2322.166	1439			

Furthermore, two aspects were additionally considered: The social strata and the highest level of education of the participants. The social strata was found not to be significant due to the homogeneity and the unequal distribution amongst the social strata. However, the finding regarding the social strata was not particularly surprising, because the social strata was affected by the choice of the sample (students) who usually stem from higher socio-economic backgrounds (as mentioned in the results section of the main experiment on page 49). Although testing the rating differences among the different levels of education amongst students did show statistically significant results for happiness and anger, this result must be treated as tentative until more research has been conducted.

Discussion

The main experiment was designed to test the hypotheses and sub-hypotheses of the study: (1) If any difference among the viewing conditions of the independent variables *stated gender* and *intended expression* was observed; (2) how strong the two aforementioned variables affect the ratings on the intensity of the different emotions, and (3) if any tendency was observed among the different groups or conditions of the independent variables *participants' biological gender*, the *participants' social gender*, and their *musical education*. Even if only a small influence was assumed, the *real biological gender of the pianists*, the *social strata*, and the *highest level of education of the participants* also were analysed.

The analysis did not support the assumption that the artists' (biological) gender affects the ratings more strongly than the level of expression. On the contrary, it was shown by the data that the reverse was true. In the absence of any differences between the genders, it can be concluded that the findings also contrast with the outcome of Behne (1990), who stated that the gender of pianists affects ratings of expressed emotions and the accuracy in piano playing more than the music does. Furthermore, the differences between rating women and men in emotion ratings discussed by Locke (2002) and Brody and Hall (2008) were not found in this study (concerning the stated gender of the pianists). However, a difference among the real genders of the pianists was observed. It was not tested how strong the influence of the real gender of the pianists on the intensity ratings was, but this could support the findings of previous researchers. Still, until more research is conducted concerning this issue, this finding must be treated as tentative. Besides findings concerning the main hypothesis, a number of conclusions may be drawn from the analysis of the observed data concerning the additional assumptions made in advance. In addition, the findings pointed to several not assumed relations and differences among some of the tested groups and conditions. However, on the evidence presented, it cannot be inferred with certainty whether the additional findings are consistently valid or not. A complete summary of the findings of the main experiment of this Master Thesis is discussed in the following paragraphs:

Stated gender & intended expression. As Figure 16 illustrates, no statistically significant difference was found among playing conditions with differently stated biological gender. This stands in contrast with the assumption that the stated gender does affect ratings in intensity of emotions. At the same time, the analysis of the intended expression playing conditions showed that a high statistically significant difference was observed for fear, happiness, and sadness, whilst anger achieved a lower statistically significant level. These findings are consistent with the assumption that the expression level affects the ratings of emotions. Moreover, it can be concluded from Figure 17 that happiness and anger tend to be associated with an exaggerated level of expression while, at the same time, fear is associated with a deadpan, and anger with a normal level of expression. More generally, it can be predicated that different emotion categories are linked to certain levels of expression. This finding is consistent with Dahl and Friberg (2004) who found that the movement cues correlate with emotional expression. The findings of Dahl and Friberg also support the association of anger and happiness with an exaggerated (or high) level of expression while fear was associated with the lowest amount of body movement (deadpan level of expression).

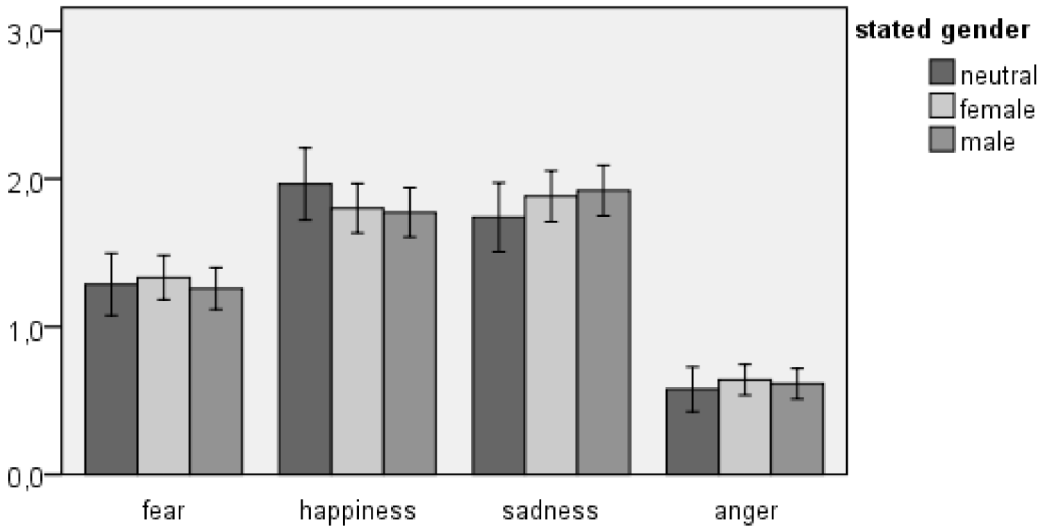


Figure 16. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the playing conditions of the stated gender 'neutral', 'female', and 'male' (error bars: 95% confidence interval).

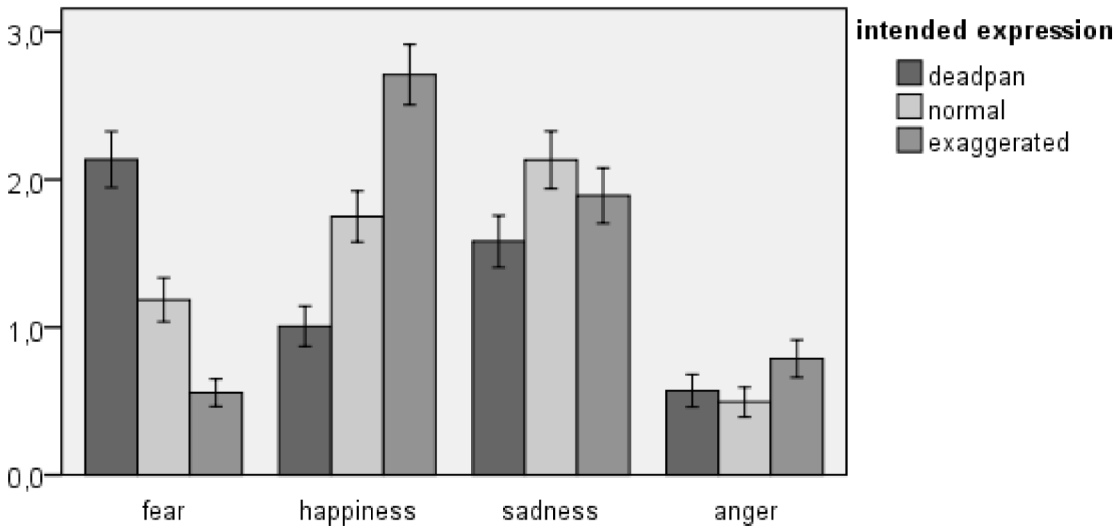


Figure 17. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the playing conditions 'deadpan', 'normal', and 'exaggerated' (error bars: 95% confidence interval).

On this basis, it may be concluded that the assumption that 'gender affects the perception of emotions more strongly than motion' cannot be validated by the findings of the experiment. On the contrary, as indicated by the figures, it can be assumed that the reverse is the case. However, this finding must be interpreted with caution: It was shown by the linear regression that the expression explains approximately 13% of the total variance for the emotion fear, and even less for the other three emotions. However, this value is high compared to the stated gender which explains around 1% and less (without achieving statistically significance).

Even if the main hypothesis was not initially confirmed by the results, they indicated that the assumption does not need to be rejected in general due to an additional analysis: As can be seen from Figure 18, the intensity ratings differed among the real genders of the performing pianists. This finding was statistically significant for all four emotions, and even achieved a high statistically significant difference for the emotions happiness, sadness and anger. Furthermore, it may be concluded, that male pianists tend to be associated with emotions expressed with an exaggerated level of expression. The figure shows that there are large variations in the mean ratings of the intensity of happiness and anger. However, this finding must be interpreted with caution because no direct relation was observed throughout this experiment.

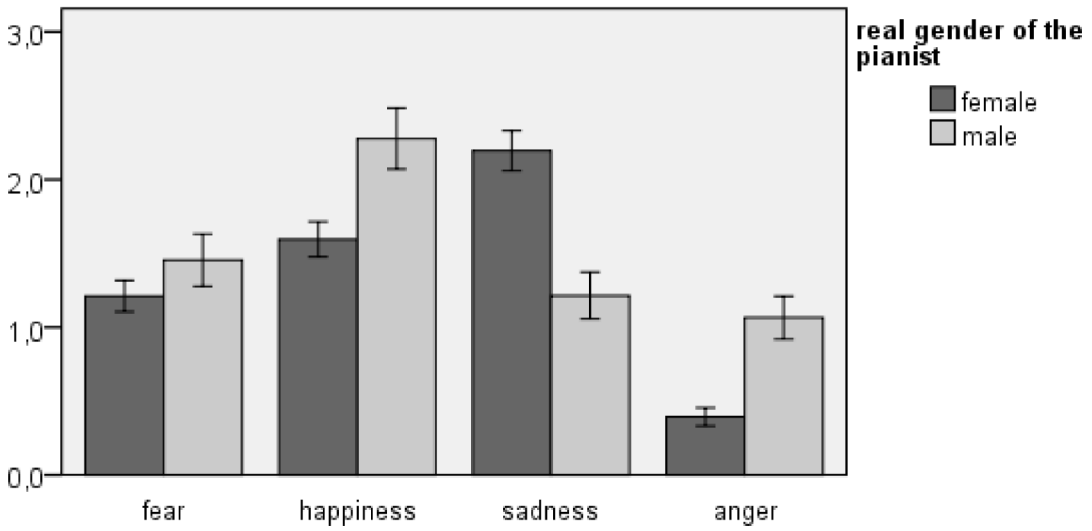


Figure 18. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the real biological gender of the pianist (error bars: 95% confidence interval).

This unexpected result might be explained due to the information provided to the perceiver through the point-lights. As discussed by Walk and Homan (1984), Dittrich et al. (1996), Dahl and Friberg (2004), and others, multiple amounts of information can be decoded by a participant even if only a point-light recording is shown to him or her. However, this explanation would include that male pianists' performances differ significantly from female pianists' performances in the expression of emotion and therefore confirm previous research (e.g. Behne, 1990, Brody & Hall, 2008). However, this result is not significant for mainly two reasons: (1) The gender of the recorded pianists was unequally distributed, and (2) only one male and two female pianists were recorded for the videos. A larger sample of different pianists would be required to reach significance and in order to draw conclusions. Still, this finding is particularly interesting with regard to previous research.

Gender categories. The assumption that individuals attributed to a different social gender perceive (rate) the intensity of emotions differently could be validated, as shown in Figure 19. Still, for the emotion 'happiness' no statistically significant difference was observed among the different gender categories. Moreover, as indicated by the figure, it can be concluded that androgynous individuals tend to rate the emotions

anger, fear, and sadness higher in general than individuals that were attributed to other gender categories, even though this difference was not found to be significant between androgynous and masculine individuals (in consideration of the error bars). Furthermore, it was found that androgynous individuals differ statistically significantly from feminine and undifferentiated individuals in intensity ratings of emotions. This finding was not observed for androgynous and masculine, genders who partly rated similarly (this was the case for the emotions fear and anger). However, due to the small sample size and the homogeneous sample (only students who usually tend to be androgynous) this finding must be interpreted with caution. Nevertheless, the results support the assumptions discussed in recent research (e.g. Wester et al., 2002, Shields et al., 2006) that the (social) gender affects the perception of emotion. Still, further research must be carried out to investigate the direction of the differences among the gender categories. Furthermore, additional tests should be used to measure the gender of the participants to compare the outcome and possibly validate the findings of this experiment. Nevertheless, the findings were highly statistically significant and therefore ought to be considered as valid.

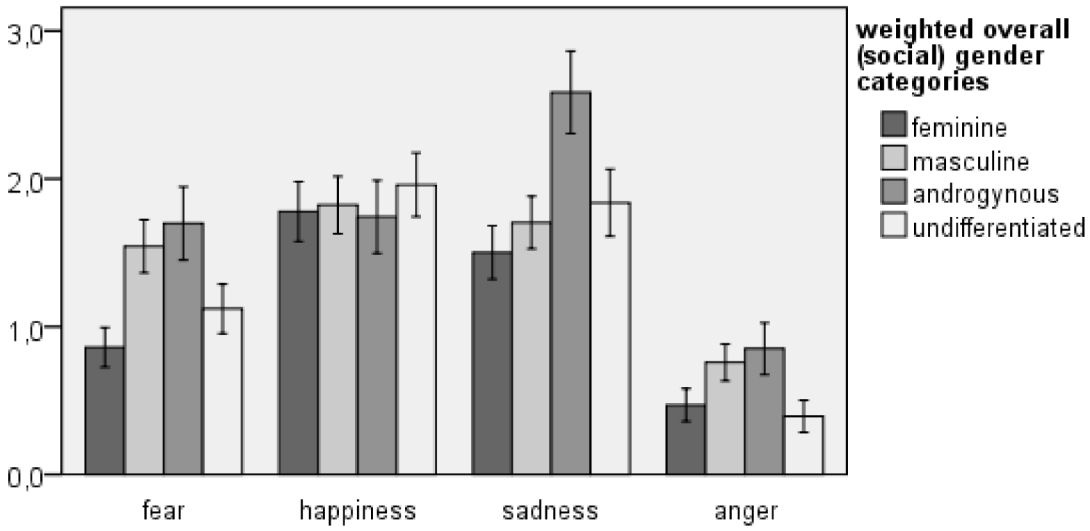


Figure 19. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the gender categories 'feminine', 'masculine', 'androgynous' and 'undifferentiated' (error bars: 95% confidence interval).

Participants' biological gender. A statistically significant difference between female and male participants was observed as can be seen from Figure 20. The findings also indicate that male individuals tend to rate emotions higher in intensity in general. This finding is partly consistent with previous research (e.g. Wester et al., 2002, Shields et al., 2006) who suggested that ratings of female and male individuals differ among each other. Interestingly, in particular the difference in intensity ratings of the emotion anger achieved a high statistically significance with the tendency of male participants rating higher than female subjects.

Comparing the results of the comparisons among the gender categories to those among the participants' biological genders, it may be concluded that the differences in gender also might be traced back to only two correlating types. On the evidence presented, such combinations could be androgynous and masculine gender on the one hand, and feminine and undifferentiated gender on the other. The differences between groups in such combinations might be similar to the ratings of male and female participants. As shown in Figure 19 and Figure 20, this correlation between male participants and/or the androgynous-masculine gender group differed significantly from the female participants and/or undifferentiated-feminine gender group for the emotions fear and anger. The findings discussed in this paragraph still have to be considered as preliminary results until more research is conducted to measure the relationship between the biological and the socio-culturally constructed (social) gender of a person in rating the intensity of emotions.

Furthermore, the comparisons also suggested that, even if expected otherwise, the biological gender of the participants still seems to be at least equally important for the intensity ratings of emotions as the gender of the participants is. It's not clear whether this statement is significant or not, but it shows that the assumed importance and the discussed shift from biological gender differences to social gender differences (see Wester et al., 2002) could not be validated in this study.

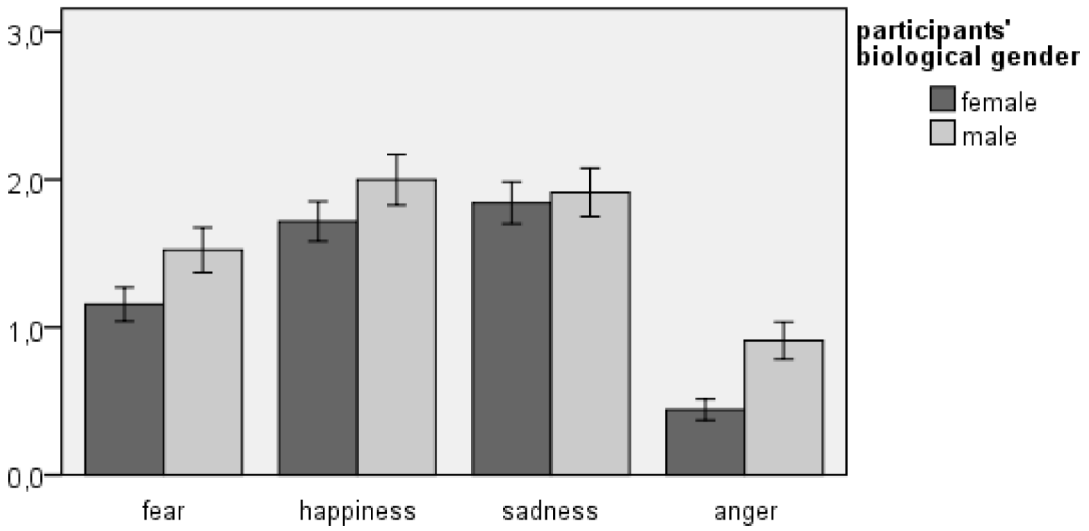


Figure 20. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the participants' biological gender (error bars: 95% confidence interval).

Figure 21. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the participants' musicality (error bars: 95% confidence interval).

Participants' background. The assumption that the musical education affects the ratings was only slightly supported for the emotion sadness, and none of the other observed emotion categories (see Figure 21). Furthermore, the social strata of the participants were too homogeneous in order to determine a significant difference in the perception of emotions. In addition, the highest level of education was observed also without any statistical relevant result.

The lack of difference between musically trained and non-musical subjects might be explained by the viewing condition: After the study, several participants commented that the videos were too abstract to be rated concerning emotionality. Therefore, it can be assumed that the reason might be the missing connection between musical performance (a real pianist) and the video clips presented to the participants. However, no obvious conclusion explaining the differences among the social strata could be drawn. Finally, the findings that no significant differences were observed for the level of education were consistent with the assumed outcome.

On the basis of the lack of correlation and/or differences among several aspects concerning the participants' educational and social background, it could be concluded that differences in ratings are not explained by any of those backgrounds. Due to the limitations because of the used sample (students only), no further conclusion was drawn from the results presented in this Master Thesis.

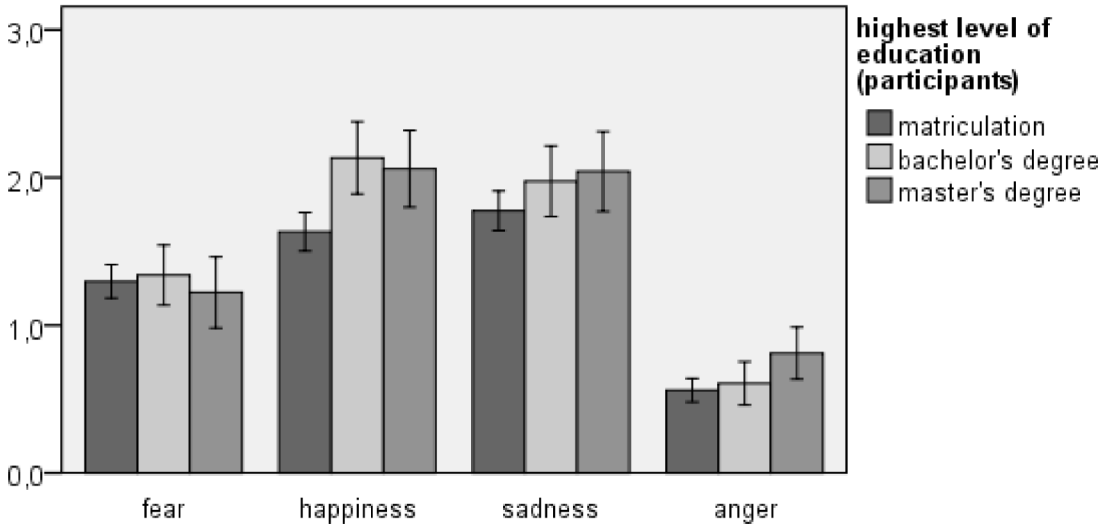


Figure 22. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the participants' highest level of education (error bars: 95% confidence interval).

Overall findings. In addition to the very specific findings among the different independent variables, two additional findings were observed across all (or most of the) variables: (1) It was observed that most participants tended to rate anger low in intensity or completely avoided to rate this emotion category in general. The low ratings in the

intensity of anger may be caused by the choice of the piece of music or by the absence of an auditive stimulus; and (2) even though the ratings for the emotion category anger were lower than the ratings for other emotions, the findings indicated that the relative difference among the conditions or groups were higher in average for anger.

In addition to the quantitative data, some of the additional comments by the participants concerning the study should be mentioned. Similar comments were combined and are presented as one comment (group) in the following. However, only the four most relevant comments were chosen to be presented in this paper. Therefore, the three far most mentioned comments are presented, and one additional finding, that is relevant because of its content:

Most of the participants (a total of 27, 84.38%) asked if the sound was not working right, if it was not part of the study to hear the audio, or similar questions. It was also asked three times which song (from German: 'Lied') they 'were seeing' in the videos. However, the participants were told that the pianists were performing a musical phrase only, and that the audio was not important. After finishing the study, the actual information about the study was provided. 13 of those participants also asked if the piece shown to them was the same each time. This last information was answered with a short 'yes' after which no further questions were asked concerning the piece of music.

A total of 23 subjects (71.88%) referred to the emotional category of emotions they '*had to choose*' in their comments. It was mentioned several times that all videos seem to be sad due to the small amount of movement.

18 participants (56.25%) mentioned that the information provided to them in the videos was very abstract, or they 'perceived everything very reduced'. Consistent among all of those participants was that in addition they admitted to rating at least one emotion lower in intensity than others (even if the emotion mentioned was not the same for all of them).

Interestingly, 4 of the participants mentioned that the names shown in the video clips affected their rating. Surprisingly, one of them additionally asked if the names in the videos were names of participants.

The comments point out that music affects which emotions and how intense they are perceived. However, this does not affect the (quantitative) results of this study, but contrasts with findings by Behne (1990) that the music itself is less important for the perception of emotions. Furthermore, the data indicates that even if stated different by the participants, the reduced information in the videos is still enough to measure differences in rating the intensity of emotions. This finding shows that a stimulus similar to the one used in this study would be appropriate to further research as well. Finally, it has to be mentioned that at least four of the participants admitted basing their ratings on the names presented to them in the videos. Still, this statement cannot be validated nor rejected, but it is considered important enough to be mentioned in this thesis.

Methodological limitations. As already mentioned in the first and second pre-experiment, one limitation was the small sample size. However, effort was made to keep the sample homogeneous across most of the possible confounding factors. Therefore, a sample with only students was chosen, all with the same cultural background to avoid misunderstandings in the experiment instructions or different meanings of the names. The significance level of $p < .001$ for several part analyses confirms that the combination of the chosen sample and the study design was still appropriate. Secondly, the self-reported data in the German Extended Personal Attributes Questionnaire (GE-PAQ) might possibly be methodologically limited as well. This influence was tried to be minimised using an adjustment for the analysis. Furthermore, since the GE-PAQ has been being developed and tested over several years by different researchers, it was therefore accepted (without expecting limitations). Last but not least, the third limitation was caused by the missing audio stimuli in the main experiment. However, in preparation of the experiment, the auditive stimulus was not found to affect the intensity ratings which was also suggested by the literature (as mentioned before).

Summary. Summarising all findings of the main experiment, the results can be divided in two main groups: Findings concerning (1) the pianists/artists, and (2) the participants: For the first group the finding did not support the hypothesis for the viewing conditions of the used stimuli (name labels), but suggested that the real (biological) gender of the pianists does support the assumption; as assumed, several differences were found among aspects concerning the participants: Social as well as biological gender differences were found to affect the perception of emotion, even though a correlation between the biological male gender, androgynous and masculine gender, as well as between biological female gender, and undifferentiated and feminine gender was suggested by the data.

Further research. Further research on this topic might include the music in the investigation. This could confirm or reject the assumed influence of the audio stimulus on the total outcome of this study. Furthermore, it was shown that the findings of this Master Thesis partly contrast with earlier findings that the gender of an artist plays an important role in the perception of emotions. Therefore, further research might explore how different aspects concerning the artist such as the intended level of expression, the gender of the artist, or even the absence of music affects the perception of emotions. Additionally, as shown in this study, research could also be conducted concerning the correlations and differences between biological and social gender differences. Moreover, further research could also focus on the tendencies discussed in this thesis. The finding that male pianists are associated with more expressive emotions such as happiness and anger would also be an issue for further research. Needless to say, as shown in this last paragraph, much additional research on this topic remains to be carried out.

6 General Discussion

Is the perception of emotion more strongly affected by the biological gender of the performer or by her/his body movement? Do other aspects, such as biological or social gender or the musical education, also affect our evaluations? This study was designed to investigate the relationships among different aspects that might affect the perception of emotional content. This general discussion of the present thesis summarises the most important aims of the study, shows various advantages and disadvantages of the general method used in the experiments (digital vs. traditional questionnaire) and discusses the most important findings.

Summary

As already mentioned in the introduction, one major aim of the study was to link three fields of research in one musicological study: (1) Biological and/or social gender differences, (2) gestures/body movement in musical performances, and (3) perception of emotions. This link was implemented through several aspects: Firstly, the research question referred to all three fields. Furthermore, in addition to the main question, an additional question referring to (social) gender differences was considered throughout the study.

Within the scope of this thesis, it was not possible to consider all related details. Therefore, several aspects (such as background of the participants, used material/videos) were restricted or simplified. The material used in the study was taken from another study and had, therefore, been tested on applicability before being used in the present experiment. An unbalanced (and small) amount of female and male pianists performing in the motion capture videos was provided for this study. Due to these restrictions, no clear conclusions concerning the real gender of the pianist could be made. Furthermore, the accuracy of recognising emotions could not be investigated. Nevertheless, the first pre-experiment showed that the videos were seen as gender neutral, and, therefore, could be used for the experiment. It can be seen from these few aspects that the study (if repeated) could be improved upon in several ways, even though the study already benefited from the numerous advantages of a digital questionnaire.

Due to several potential advantages of the use of a digital, computer-based questionnaire, several parts were conducted using the software PureData: The first pre-experiment was partly conducted on a computer – a PureData patch was used to perform the main part of the experiment, whilst the demographic part and the feedback questionnaire were performed in a *traditional method*[72]. For the second pre-experiment, a *traditional* print version of the questionnaire was used. The complete main experiment was conducted using digital questionnaires – for the main part, the gender questionnaire, as well as the social strata part, and the demographic part.

Pros and cons of the use of a computer-aided study design. The following paragraphs deal with the aforementioned pros of the use of digital computer-aided study designs but also the cons (compared to the traditional method using printed versions):

The most obvious advantage of the computer-aided tests over traditional tests might be that no manual input of data is needed. In the present study, for example, all data was available as plain text files (*.txt) or as comma-separated value file (*.csv) and could be imported directly into any spreadsheet application (such as Microsoft Excel, or LibreOffice Calc) or software for statistical analysis (such as SPSS, PSPP, or R). No matter how large the sample size, the data would be immediately available. Therefore, this can be seen as a time-related advantage. Furthermore, data input mistakes would be reduced.

Another significant advantage is the possible control of the sequence of experimental events and the dynamic embedding of several elements into the graphical interface. Both possibilities were used in this study: On the one hand, the videos were started and stopped automatically, on the other, the question was slightly adjusted for each of the videos (change of the name). Furthermore, this could also avoid possible bias caused by the investigator (for example, disturbing the test sequence while starting and stopping the video). Moreover, it was possible to ask supplementary questions without using hybrid questions (such as 'If yes, how long have you been playing an instrument?' as a supplementary question to 'Are you playing any instrument?'). This could be realised due to the possible dynamic change of questions and sections of the

72 The term *traditional* is used for a printed (non-digital) questionnaire.

graphical interface. These additional questions were hidden in the interface of the computer-aided questionnaire, and only shown if, for example, the option *yes* is selected in the main question (as shown in Figure C-1.3, appendix C, page 193 – a screenshot of the demographic sub-patch).

Furthermore, the experiment could be repeated as often as required without preparing additional material (for example, printing additional questionnaires), and even more importantly, the data input time was (almost[73]) the same for any number of participants assuming a digital questionnaire would be used. Moreover, the digital questionnaires could be seen as environmentally friendly, since no paper would be used, and therefore also is lower in cost.

Besides the already mentioned advantages, a possibly lengthy decipherment of comments and feedback (due to barely legible handwritings) on the questionnaire or answers to open questions would be eliminated by using a computer-aided input. In addition, it was possible to avoid missing values and/or double answers. Both options were used in this study even if they did not guarantee that the values were valid or not.

Even if the advantages of the digital questionnaires clearly outweighed its disadvantages, still some noteworthy disadvantages remained: The traditional method might have been more acceptable compared to the digital version (this disadvantage might also have depended on the average age of the participants and, therefore, only have played a minor role in the present study). Moreover, the development of a digital questionnaire required a longer preparation time than preparing a traditional questionnaire (this disadvantage would be less pronounced if the questionnaire would be used for a large sample size). In addition, two disadvantages concerning technical issues were observed (and solved) throughout this study: (1) The graphical interface showed visual differences between differing operating systems, and (2) importing the raw data into SPSS was complicated due to the different country-specific sign for decimal separation (PureData supports the dot as decimal separation). Furthermore, to carry out the study, it was required that the same PureData version (with all additional libraries and extensions) was installed on all computers used throughout the study.

73 The only difference in data input time for computer-aided questionnaires is the computing time, whereby differences in computing time are only significant with very large quantities of data.

Hypotheses

Is emotion more strongly affected by gender or motion? This question was the starting point of the present thesis. Based on previous research, it was hypothesised that the communication of emotions would be more strongly affected by the biological gender of the artist than by his or her bodily expression. Furthermore, it was assumed that female performers would be rated higher in the intensity of sad, happy, and fearful emotions than male performers, and that for angry emotions, male performers would be rated higher in intensity. In addition to these hypotheses, a difference in ratings from participants with a different social gender (feminine, masculine, androgynous, or undifferentiated) was assumed. To investigate these hypotheses, two additional pre-experiments had to be conducted in advance:

The first pre-experiment was used to test the motion capture videos on their applicability for the main experiment. Therefore, it was hypothesised that the videos were seen as gender neutral (neither male nor female). In addition, it was assumed that a text label in the bottom right corner would affect the reactions of the participants. Furthermore, the digital computer-aided questionnaire was tested in handling and usability in the first pre-experiment. The second pre-experiment was used as a decision support only – therefore no hypotheses concerning the results were formulated.

Conclusion

The overall conclusion of the thesis is structured into three sub-sections: (1) The additional discussion and comparison of the results of the present study and previous research, (2) implications of the findings, and (3) suggestions on improving the study and further research. This conclusion section can be seen as additional to the discussion sections of each of the three main sections, and links both pre-experiments with the main experiment. The results of the main experiment are compared with previous research with regard to the findings of the pre-experiments. In addition to the research suggested in the discussion of the main experiment (see page 75) a further research section is added. The suggested research mentioned in this chapter is strongly connected to the present study and can be conducted using a study design similar to the one used throughout this study.

Discussion and comparison of the results with previous research. A comparison of the results with previous research can only be seen as partly valid due to several limitations of each study: First of all, the present study was limited to university students of three universities in Graz (Austria), which leads to a limitation of education level, cultural background, and partly the social strata (as preliminary discussed in the results section of the main experiment, page 47). However, on the one hand, the results of this thesis suggest that the biological gender hardly affects the perception of emotion, which is in contrast with earlier findings discussed by Locke (2002), and Brody and Hall (2008). On the other hand, it has been shown that the real (biological) gender of the pianists might affect the evaluations of the participants, which would support the results of the aforementioned studies. Hence, it can be concluded that the gender differences in perception of emotions are not only gender-dependent but more dependent on how this information is provided to the perceiver.

However, previous research mainly used ordinary video recordings to investigate biological and social gender differences in perception and/or expression of emotion instead of using abstract motion capture video clips. It was found that at least some of the participants were confused by the abstraction of the human body shape used in this study. This could partly explain the differences between the results of earlier studies and the present study. However, future research could use techniques already used by the film industry to compute a female or male body shape around the point-light recordings to avoid such confusion.

The second important finding of the present study is the influence of the participants' biological and social gender on their evaluations. As discussed by Wester et al. (2000), a shift from biological to social gender, concerning behavioural and perceptual issues, might take place in the western society. However, this study did not investigate if the social and the biological gender were correlating for the participants. Therefore, it is not possible to determine whether this is true or not, however, it was found that the social gender correlates with the biological gender for the perception of emotions. This finding is not surprising, because strong female and male attributes were used to figure out the (self-rated) social gender of each participant. A concept based on

these attributes (combined to scales) was used to divide the participants in feminine, masculine, androgynous, and undifferentiated (see also results section of the main experiment, page 47, as well as Athenstadt and Altstötter-Gleich, n.d.). It can be seen in this concept that a distinction of the social gender is still based on prejudices against women and men. To avoid such stereotyped thinking, a completely new gender role system would have to be introduced. In this thesis, two terms (suggested by earlier literature and also used by Athenstaedt, 2003) were used to replace feminine and masculine in describing attributes: *Instrumental* and *expressive*. In addition, future studies might attempt to replace the terms used to describe the social gender (feminine, masculine, androgynous, and undifferentiated) with similar terms already used to describe the attributes. Nevertheless, this study validated that, for now, the terms still might be valid due to the similarities of, for example, ratings by participants with feminine gender compared to participants with female gender. This finding is also in agreement with previous studies (e.g. Wester et al., 2002, Shields et al., 2006).

Implications of the study. It has been shown that a name labelling barely affects the perception of emotions, if at all. On the other hand, on the evidence presented, it can be assumed that the real (biological) gender of a performer affects our evaluations even if the body is disguised. Therefore, a strong relation between gender and communication of emotions can be assumed even though several other possible aspects have been found to affect this communication. However, the strong relationship between bodily expression and gender of the performer and emotional expression can be used to influence the recipients perception in musical performances. This phenomenon can, for example, be used in the music industry for marketing purposes of various musical pieces by selecting certain artists (depending on their biological gender) for performances. Furthermore, music therapy might also benefit from the results: Even if the findings were not uniform across the stated and the real (biological) gender of the pianists, it was shown that female and male artists are perceived emotionally different. However, further research must be conducted to explain the exact influence of the factor of gender in various forms of music therapy.

Suggestions for improving the present study and further research. One major limitation of all three study parts (first pre-experiment, second pre-experiment, and main experiment) was the small amount of participants. Therefore, to improve the study (if repeated) an online questionnaire is suggested (especially for the second pre-experiment). This methodological change would increase the number of participants, lead to a better distribution of the sample (and therefore better representativeness), as well as give the possibility to consider differences among different groups, if any. However, the results of such online investigations would need to be interpreted with caution[74]. Furthermore, the use of a larger amount of motion capture video clips, as well as an equal amount of female and male pianists could help to enhance the validity of the study. Two optional suggestions for improving or changing the study design only relate to the second pre-experiment: Either the scale could be reduced to a 3-step scale (female – neutral – male) to force the participants to make a decision, or the actual method could be replaced with a free association. However, both suggested methods have their limitations. In addition to these improvements, other options for labelling the videos could be tested (for example, showing female, male, or gender neutral connoted pictures/faces to the participants).

As discussed in the preceding paragraph, the whole study could be improved in several ways. For example, a repetition of the present study with the suggested methodological improvements could therefore be one possible future study. Moreover, this study could be repeated with instruments other than piano (or keyboard instruments in general) and be compared to the outcome of the present study, for example, with regard to similarities in perception of special movement cues. The study design could also be used to build on previous research and consider how strong the real (biological) gender of the performer affects the perception of emotions compared to body movements.

74 The use of online questionnaire also can lead to misleading results: Multiple responses by several participants, no control of the sample, no control of the experimental situation.

This thesis suggests that, besides bodily expression, several aspects (such as biological and social gender differences) affect the perception of emotion in musical performances. One very interesting finding was that the biological and the social gender contain parallels in ratings. However, the question to what extent these aspects affect the perception cannot be answered without further research. Furthermore, the gender of the perceiver – biological and social gender – was observed to affect the evaluations on intensity of different emotions.

In conclusion, the findings of the study are shortly summarised: To conclude, body movement was observed to affect the perception of emotion more strongly than gender. Furthermore, it was shown that male pianists tended to be rated higher in the intensity of emotions expressed with an exaggerated level of expression (happiness and anger), whilst female pianists were associated with lower levels of expression (in particular sadness). The evaluations were also affected by the gender of the perceiver.

7 References

Altstötter-Gleich, C. (2004). Expressivität, Instrumentalität und psychische Gesundheit: Ein Beitrag zur Validierung einer Skala zur Erfassung des geschlechtsrollenbezogenen Selbstkonzepts. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *25*(3), 123-139.

Athenstaedt, U. (2003). On the content and structure of the gender role self concept: Including gender-stereotypical behaviours in addition to traits. *Psychology of Women Quarterly*, *27*(4), 309-318.

Athenstaedt, U., & Altstötter-Gleich, C. (n.d.). *The Four Facets of the Gender Role Attribute Self-Concept.* Unpublished manuscript, University of Graz, Austria.

Aviezer, H., Hassin, R. R, Ryan, J., Grady, C., Susskind, J. Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? *Psychological Science*, *19*(7), 724-732.

Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition and Emotion*, *12*(4), 579-599.

Barrett, L. F. (2011). Was Darwin wrong about emotional expressions? *Current Directions in Psychological Science*, *20*(6), 400-406.

Behne, K.-E. (1990). “Blicken Sie auf die Pianisten?!” Zur bildbeeinflußten Beurteilung von Klaviermusik im Fernsehen. *Medienpsychologie*, *2*(2), 115-131.

Brody, L. R., & Hall, J. A. (2008). Gender and emotion in context, in: M. Lewis, J. M. Haviland-Jones, L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 395-408), New York: Guilford Press.

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion*, *24*(3), 377-400.

Bundesministerium für Wissenschaft und Forschung (Ed.) (2010). *Materalien zur sozialen Lage der Studierenden 2010*. Wien: BMWF.

Bundesministerium für Wissenschaft und Forschung (Ed.) (2012). *Materalien zur sozialen Lage der Studierenden 2012*. Wien: BMWF.

Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., & Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, *34*, 1171-1180.

Clore, G. L., & Ortony, A. (2008). Appraisal theories: How cognition shapes affect into emotion, in: M. Lewis, J. M. Haviland-Jones, L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 628-644), New York: Guilford Press.

Dahl, S., & Friberg, A. (2004). Expressiveness of musician's body movements in performances on marimba, in: A. Camurri, & G. Volpe (Eds.), *Gesture-Based Communication in Human-Computer Interaction: Vol. 2915. Lecture Notes in Artificial Intelligence* (pp. 479-486), Berlin: Springer.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in music performance, *Music Perception*, *24*(5), 433-454.

Darwin, C. (1965). *The expression of the emotions in man and animals*. Chicago: University of Chicago Press. (Original work published 1872)

Davidson, J. W., & Correia, J. S. (2002). Body movement, in:R. Parncutt, & G. E. McPherson (Eds.), *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning* (pp. 237-250), New York: Oxford University Press.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*(6), 727-738.

Ekman, P. (1999). Basic emotions, in: T. Dagleish, & M. J., Power (Eds.), *Handbook of Cognition and Emotion* (pp. 45-60), Chichester, England: John Wiley & Sons.

Ekman, P. (2003). *Emotions Revealed*. New York: Times Books.

Ellis, L. (Ed.). (2008). *Sex Differences: Summarizing more than a century of scientific research*. New York: Hove.

Fischer, A. H., Rodriguez Mosquera, P. M., van Vianen, A. E., & Manstead, A. S. (2004). Gender and culture differences in emotion. *Emotion*, *4*(1), 87-94.

Gosselin, P., Kirouac, G, & Doré, F. Y. (2005). Components and recognition of facial expression in the communication of emotion by actors, in: P. Ekman, & E. L. Rosenberg (Eds.), *What the Face Reveals: Basic and Applied studies of Spontaneous Expression using the Facial Action Coding System FACS* (2nd ed., pp. 243-267), New York: Oxford University Press.

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*(4), 845-857.

Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive stlye*. Baltimore: Johns Hopkins University Press.

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

Hall, J.A., Carter, J. D., & Horgan, T. G. (2000). Gender differences in nonverbal communication of emotion, in: A. H. Fischer (Ed.), *Gender and Emotion: Social Psychological Perspectives* (pp. 97-117), Cambridge: Cambridge University Press.

Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation? *Personal Relationships*, *7*(1), 95-109.

Juslin, P. N., Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770-814.

Kreutziger-Herr, A., & Unseld, M. (Eds.). (2010). *Lexikon Musik und Gender*. Stuttgart: Metzler.

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. *Emotion*, *6*(1), 125-138.

Locke, A. (2002). Gendered emotion: Personal, cultural or discursive? *Feminism & Psychology*, *12*(1), 97-104.

Matsumoto, D., Franklin, B., Choi, J.-W., Rogers, D., & Tatani, H. (2002). Cultural influences on the expression and perception of emotion, in: W. B. Gudykunst, & B. Mody (Eds.), *Handbook of International and Intercultural Communication* (2nd ed., pp. 107-126), London: Sage Publications.

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *PNAS (Proceedings of the National Academy of Science of the United States of America)*, *102*(45), 16518-16523.

Milovchevich, D., Howells, K., Drew, N., & Day, A. (2001). Sex and gender role differences in anger: An Australian community study. *Personality and Individual Differences*, *31*(2), 117-127.

Montepare, J., Koff, E., Zaitchik, D., & Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behaviour*, *23*(2), 133-152.

Niedenthal, P. M. (2008). Emotion concepts, in: M. Lewis, J. M. Haviland-Jones, L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 587-600), New York: Guilford Press.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*(3), 315-331.

Pollick, F. E., Paterson, H.M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51-B61.

Runge, T. E., Frey, D., Gollwitzer, P. M., Helmreich, R. L., & Spence, J. T. (1981). Masculine (instrumental) and feminine (expressive) traits: A comparison between students in the United States and West Germany. *Journal of Cross-Cultural Psychology*, *12*(2), 142-162.

Resnicow, J. E., Salovey, P., & Repp, B. H. (2004). Is recognition of emotion in music performance an aspect of emotional intelligence? *Music Perception*, *22*(1), 145-158.

Rosip, J. C., & Hall, J. A. (2004). Knowledge of nonverbal cues, gender, and nonverbal decoding accuracy. *Journal of Nonverbal Behaviour*, *28*(4), 267-286.

Saragovi, C., Aubé, J., Koestner, R., & Zuroff, D. (2002). Traits, motives, and depressive styles as reflections of agency and communion. *Personality and Social Psychology Bulletin*, *28*(5), 563-577.

Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395-399.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*(6), 1061-1086.

Shell Deutschland Holding (Ed.) (2010). Appendix 'Dokumentation des Index der Sozialen Schicht' of *16. Shell Jugendstudie: Jugend 2010: Eine pragmatische Generation behauptet sich* (pp. 400-401). Frankfurt a.M.: Fischer Taschenbuch.

Shields, S. A., Garner, D. N., Di Leone, B., & Hadley, A. M. (2006). Gender and emotion, in: J. E. Stets, & J. H. Turner (Eds.), *Handbook of the Sociology of Emotion* (pp. 63-86), New York: Springer.

Tabei, K.-I., & Tanaka, A. (2012). Multisensory perception of six basic emotions in music, in: E. Cambouropoulus, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*. Thessaloniki, Greece: Aristotle University of Thessaloniki , 969-970.

Thamm, R. A. (2006). The classification of emotions, in: J. E. Stets, & J. H. Turner (Eds.), *Handbook of the Sociology of Emotion* (pp. 11-37), New York: Springer.

Thompson, M. R., & Luck, G. (2008). Effect of pianists' expressive intention on amount and type of body movement, in: S.W. Yi (Ed.), *Proceedings of the 10th International Conference on Music Perception and Cognition*. Sapporo, Japan: University of Hokkaido. 540-544.

Walk, R. D., & Homan, C.P. (1984). Emotion and dance in dynamic light displays. *Bulletin of the Psychonomic Society*, *22*(5), 437-440.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, *28*(6), 879-896.

Wester, S. R., Vogel, D. L., Pressly, P. K., & Heesacker, M. (2002). Sex differences in emotion: A critical review of literature and implications for counseling psychology. *The Counseling Psychologist*, *30*(4), 630-652.

Additional References

The following sources were not directly cited or referred to in the paper, but used to set up different parts of the study:

Beliebteste Vornamen. (n.d.). Retrieved March 5, 2012, from Vornamen Lexikon: http://www.vornamen-weltweit.de

Beliebte Vornamen. (n.d.). Retrieved March 5, 2012, from beliebte-Vornamen.de: http://www.beliebte-vornamen.de

Boys names/Girls names/Unisex names. (n.d.). Retrieved March 5, 2012, from Babynames UK: http://www.babynames.co.uk

Die Namensdatenbank. (n.d.). Retrieved March 5, 2012, from Namepedia: http://www.namepedia.org

Hahr, M. (2012). List Randomizer [Online "Program"]. Accessed April 17/18, 2012, from Random.org: http://www.random.org/lists/

Vornamen Hitparade 2010 deutsche Schweiz. (n.d.). Retrieved March 5, 2012, from Vornamen.ch: http://www.vornamen-online.ch

Open Source. (2012). PureData extended [Software/Visual Programming Language]. Retrieved from PD Community Site, a contribution of the Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz: http://puredata.info/downloads/pure-data/releases/0.43.1

8 Appendices

Appendix A0 [Pre-Experiment I – Instructions]

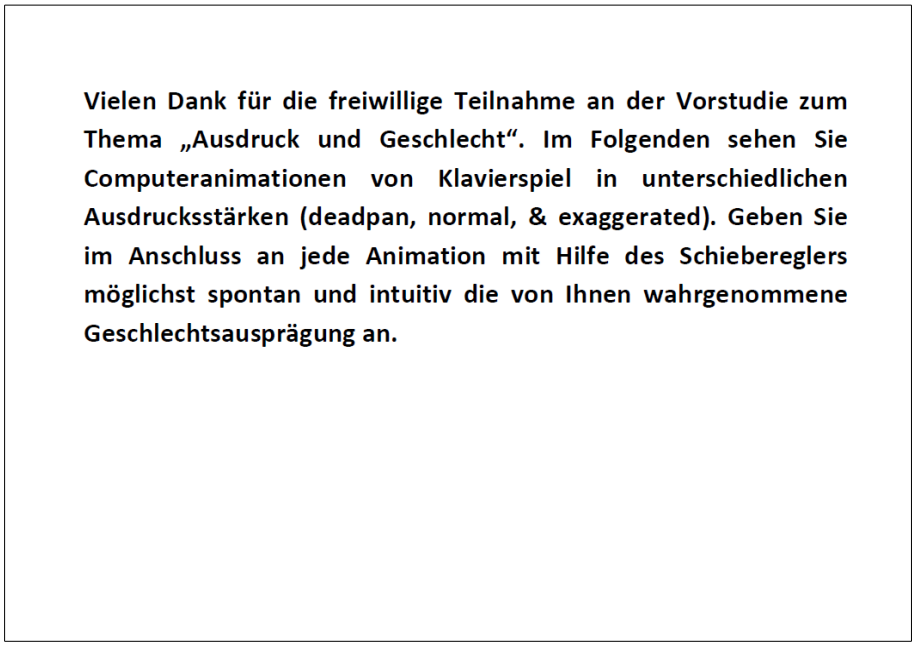


Figure A-0.1. Screenshot of the Adobe PDF file of the experiment instructions.

Original Adobe PDF available on appendix CD:
[/content/experiments/pre-experiment_#1/instructions.pdf]

Own translation of the German instructions:
Thank you for participate voluntarily in the pre-study on “expression and gender”. You will see computer animations showing piano play in different expression levels (deadpan, normal, & exaggerated) in the following. Please rate as spontaneously and intuitively as possible the perceived gender of the pianist using the horizontal slider.

Appendix A1 [Pre-Experiment I – PureData Patch]

The PureData patch was designed specifically for the purpose of this study. The patch used in the first pre-experiment is structured into several sub-patches: (1) A graphical interface, as shown in Figure A-1.1, (2) a control patch, shown in Figure A-1.2, and (3) several sub-patches used for executing the main patch and to write the results into text files (see Figures A-1.3 and A-1.4). Furthermore, a GEM window is used to open the video clips (Figures A-1.5 to A-1.13).

The patch had to be reset and started manually (see reset and start button in the graphical interface) by the investigator before starting the experiment to clear all previously stored values. Other functions (such as open and start the video clips, or reset the slider to position zero) are processed automatically and, finally, the writing process is partly automated and has to be started by clicking the write button of the control sub-patch after each experiment manually to avoid data loss.

The patch uses several built-in objects of PureData: The most important objects are a pseudo random generator to randomise the video playing order, delays to put processes in the right order (for example, opening a video after starting a GEM window), and an autoplay for GEM. Detailed information about all built-in objects used in this patch can be found in the help file of the software.

Note. The PureData patch does NOT entirely work without the video files. The videos were exclusively provided for the purpose of this master thesis and, therefore, not included in the appendix or the digital appendix. If the videos are needed, please contact Marc Thompson and ask for permission (if granted, the adjusted video clips can be provided by me).

Screenshots of the PureData Patch

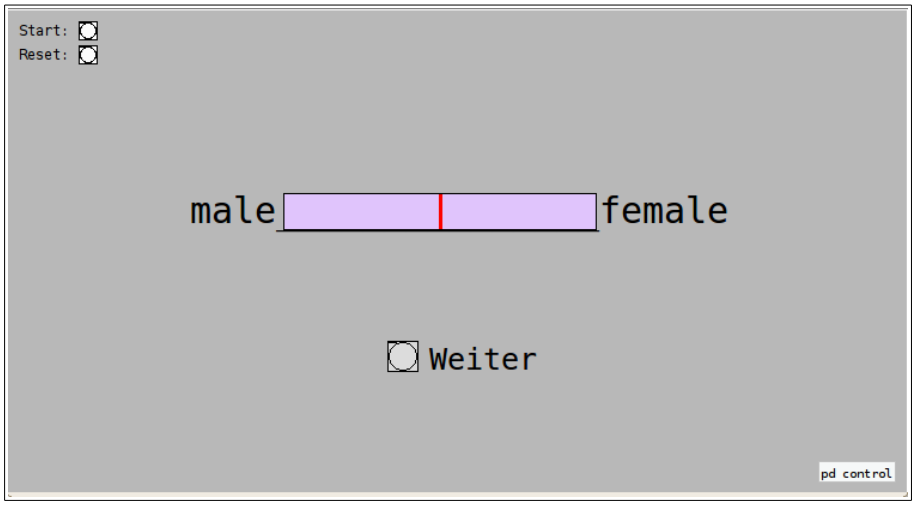


Figure A-1.1. Screenshot showing the input screen of the used PureData patch.

PureData patch available on appendix CD:
[/content/experiments/pre-experiment_#1/pre-experiment_#1.pd]

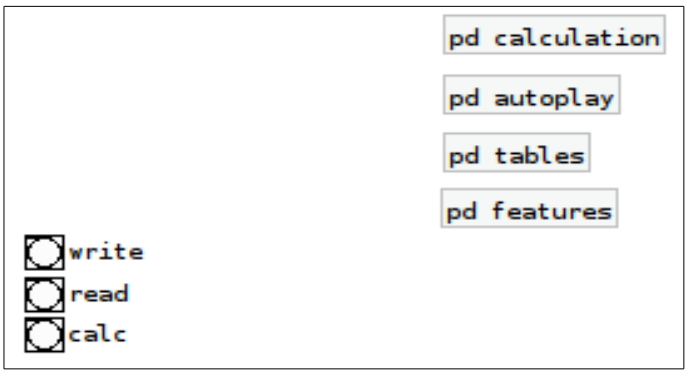


Figure A-1.2. Screenshot of the control sub-patch.

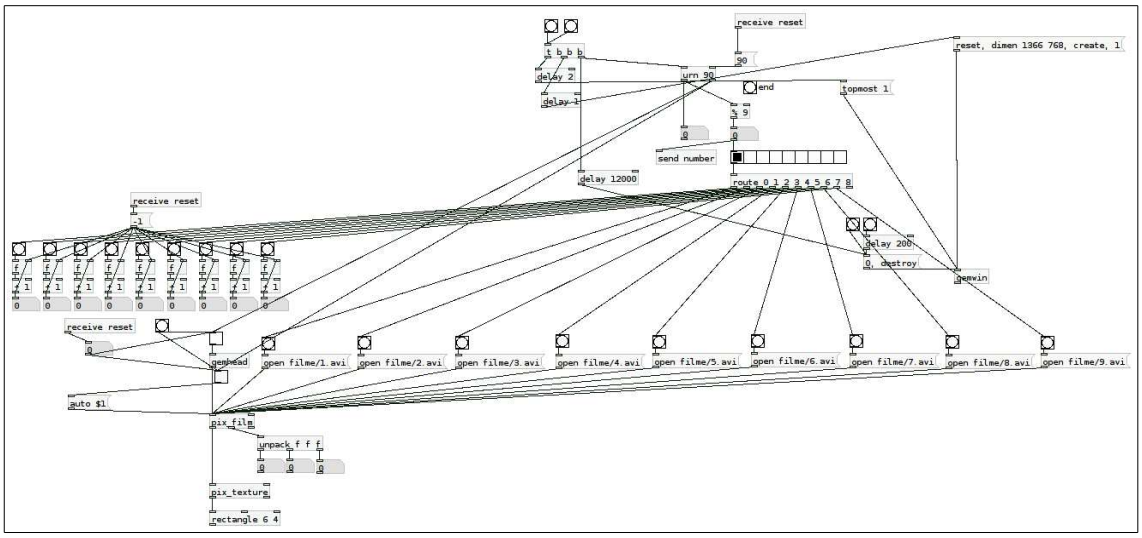


Figure A-1.3. Screenshot of the autoplay sub-patch.

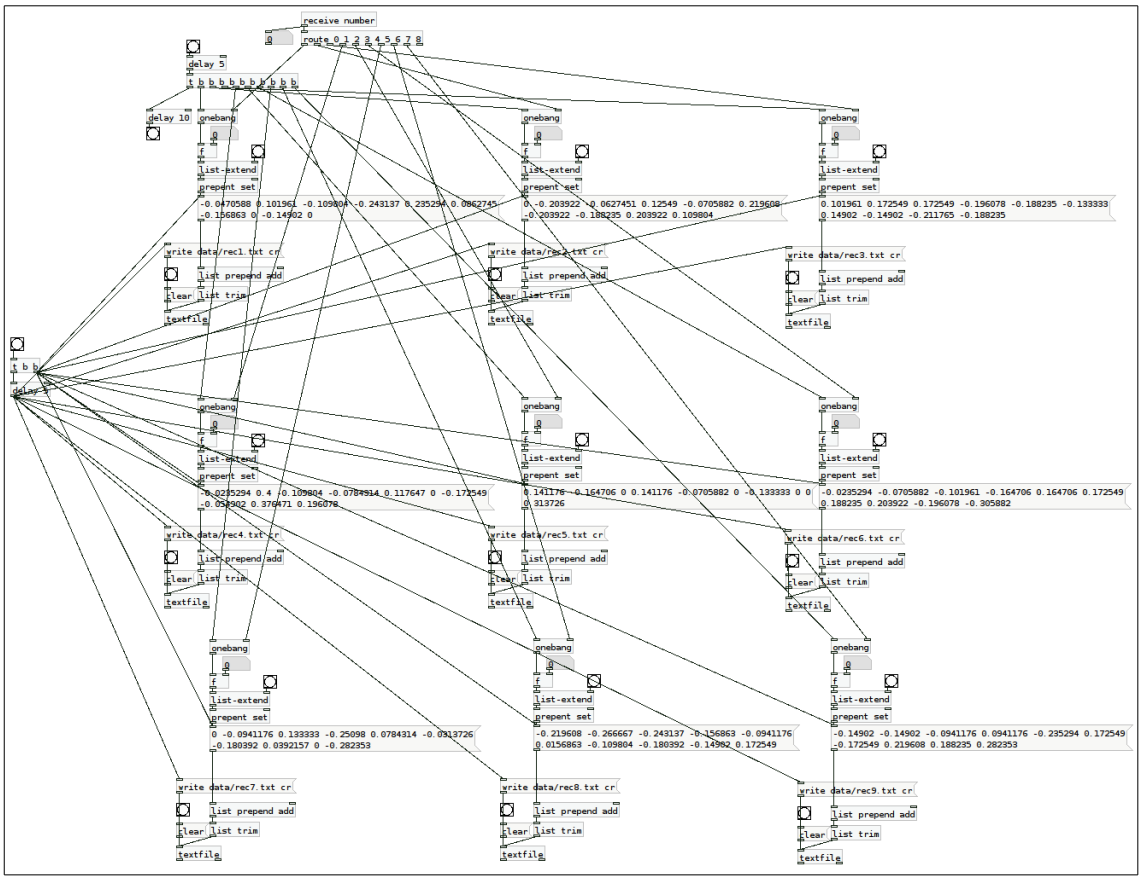


Figure A-1.4. Screenshot of the data writing sub-patch.

Video Frames

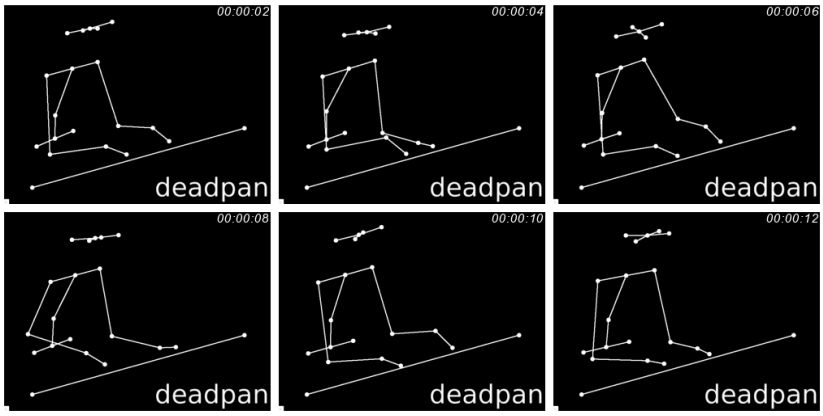


Figure A-1.5. Frames of the video no.1, originally played by a female pianist.

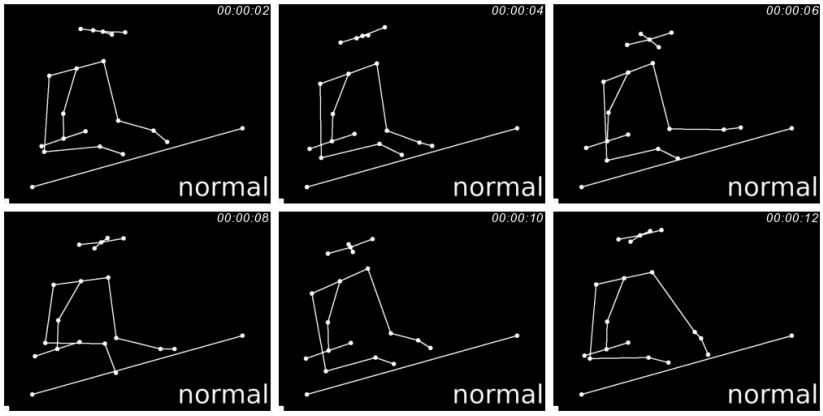


Figure A-1.6. Frames of the video no.2, originally played by a female pianist.

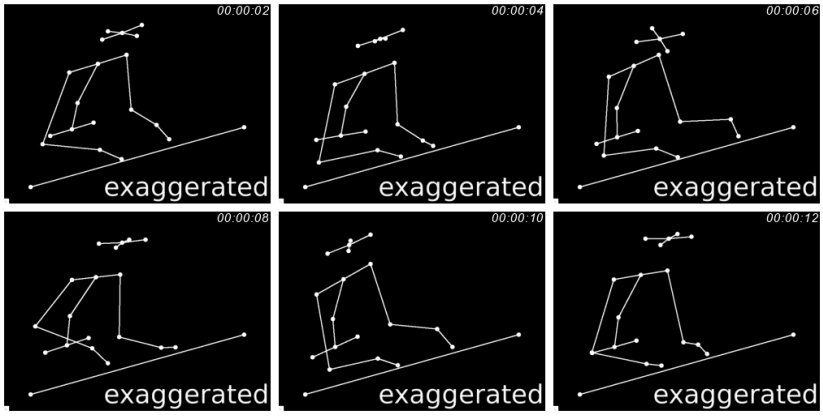


Figure A-1.7. Frames of the video no.3, originally played by a female pianist.

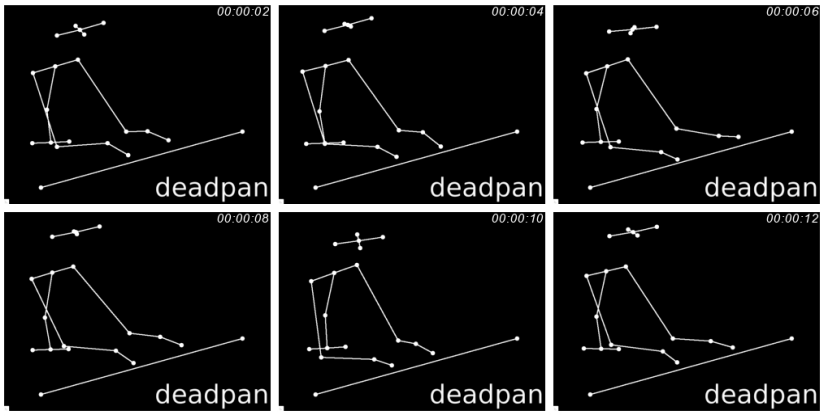


Figure A-1.8. Frames of the video no.4, originally played by a female pianist.

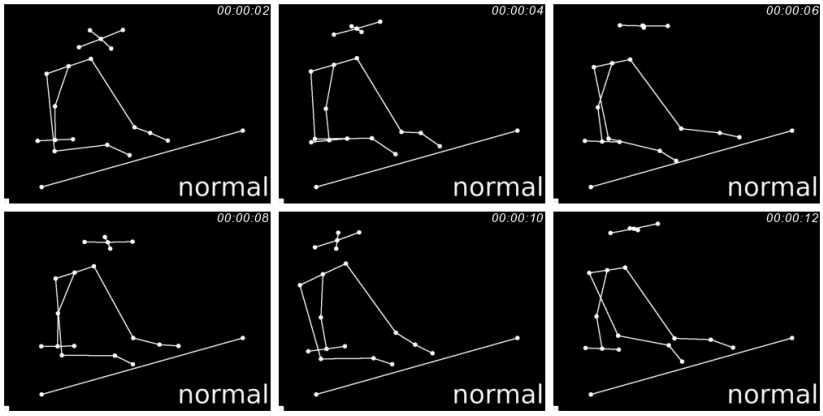


Figure A-1.9. Frames of the video no.5, originally played by a female pianist.

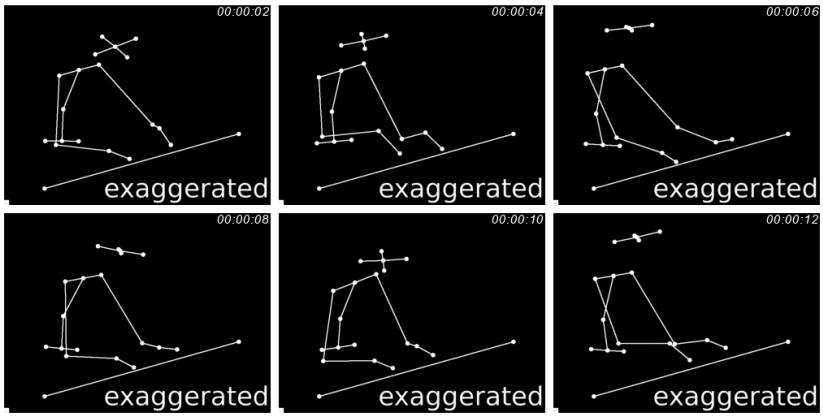


Figure A-1.10. Frames of the video no.6, originally played by a female pianist.

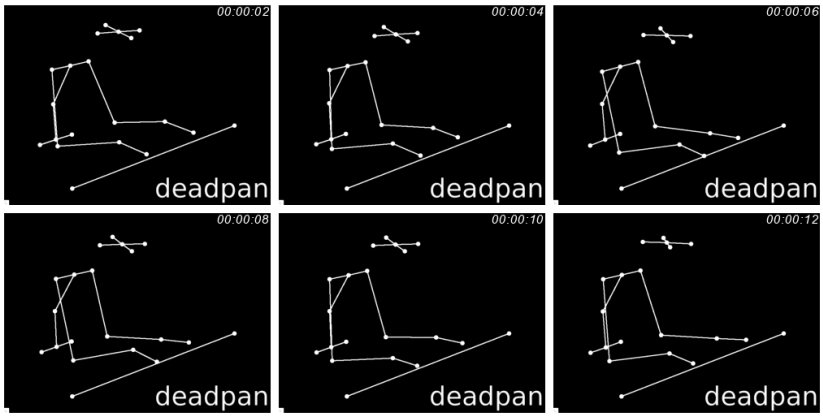


Figure A-1.11. Frames of the video no.7, originally played by a male pianist.

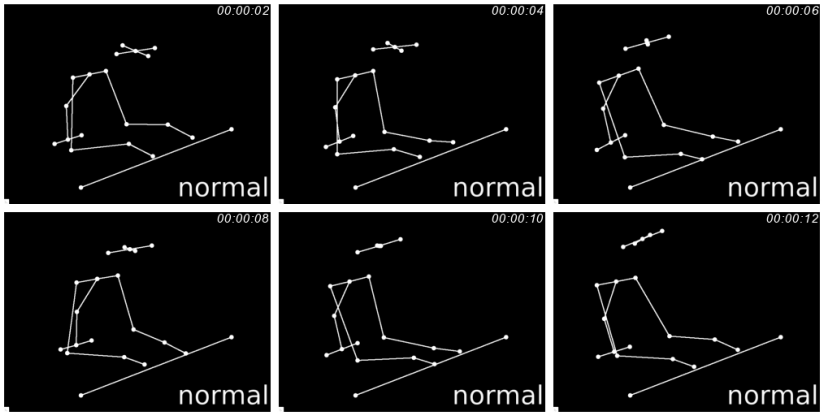


Figure A-1.12. Frames of the video no.8, originally played by a male pianist.

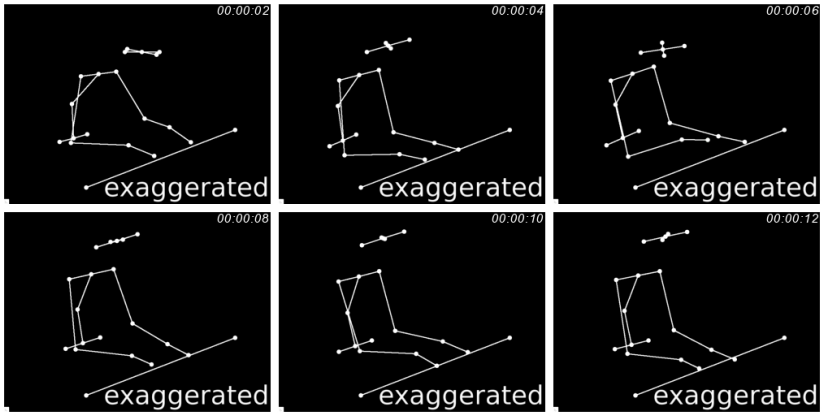


Figure A-1.13. Frames of the video no.9, originally played by a male pianist.

Appendix A2 [Pre-Experiment I – Feedback & Demographic]

Ihr Feedback ist wichtig. Beantworten Sie daher bitte die folgenden Fragen:

War die Instruktion klar und verständlich formuliert? o ja o nein

War die Instruktion zu lang? o ja o nein

War die Studie insgesamt zu lang? o ja o nein

Wie ist es Ihnen bei dieser Studie gegangen?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Geben Sie bitte noch folgende Informationen zu Ihrer Person an:

Geschlecht: o männlich o weiblich

Alter: _____

Höchste abgeschlossene Ausbildung:
o Pflichtschule
o weiterführende Schule (ohne Matura)
o Lehre
o Matura
o FH-Studium
o Universitätsstudium

Figure A-2.1. Feedback questionnaire (top) with added demographic part (bottom); the bold headings separate the parts from each other.

Original Adobe PDF available on appendix CD:
[/content/experiments/pre-experiment_#1/feedback_questionnaire.pdf]

Appendix A3 [Pre-Experiment I – Raw Data]

The values are given precisely to three decimal places. More exact values are available in SPSS file format [*.sav], MS Excel 97/2003 format [*.xls], LibreOffice Calc (version 4.0) format [*.ods], or as plain text files [*.txt] on the appendix CD [/content/experiments/pre-experiment_#1/data/..].

Table A-3.1

Raw data of subject no.2. The participant was 26 years old, student, and has not specified her or his gender. Duration of the experiment was 35 minutes.

chronology	video no.1	video no.2	video no.3	video no.4	video no.5	video no.6	video no.7	video no.8	video no.9
1st	-.047	.000	.102	-.024	.141	-.024	.000	-.220	-.149
2nd	.102	-.204	.173	.400	-.165	-.071	-.094	-.267	-.149
3rd	-.110	-.063	.173	-.110	.000	-.102	.133	-.243	-.094
4th	-.243	.125	-.196	-.078	.141	-.165	-.251	-.157	.094
5th	.235	-.071	-.188	.118	-.071	.165	.078	-.094	-.235
6th	.086	.220	-.133	.000	.000	.173	-.031	.016	.173
7th	-.157	-.204	.149	-.173	-.133	.188	-.180	-.110	-.173
8th	.000	-.188	-.149	-.055	.000	.204	.039	-.180	.220
9th	-.149	.204	-.212	.376	.000	-.196	.000	-.149	.188
10th	.000	.110	-.188	.196	.314	-.306	-.282	.173	.282

Note. The videos were played in random order, therefore the chronological order is an inner-item value only, and no impact to the total outcome was assumed.

Table A-3.2

Raw data of subject no.2. The participant was 54 years old, no student, and male. Duration of the experiment was 31 minutes.

chronology	video no.1	video no.2	video no.3	video no.4	video no.5	video no.6	video no.7	video no.8	video no.9
1st	.000	.341	.231	.302	.169	.000	-.286	.000	.263
2nd	.106	-.224	.435	-.318	.239	.247	-.443	-.239	-.616
3rd	.318	.239	.404	.224	-.129	.286	-.082	-.255	-.294
4th	.694	.451	-.435	.231	-.608	.000	-.145	-.129	.153
5th	.294	-.239	-.208	-.114	-.741	.106	-.169	-.239	-.631
6th	-.192	.412	.263	-.404	-.702	.561	-.145	.561	.576
7th	.169	.420	.224	.161	-.357	.169	-.553	.263	.514
8th	.310	.129	.153	-.161	-.624	-.396	-.608	-.145	.231
9th	.498	.357	.467	.098	-.271	-.655	-.412	-.490	.373
10th	.639	.569	.584	.349	.349	.286	-.278	-.459	.255

Note. The videos were played in random order, therefore the chronological order is an inner-item value only, and no impact to the total outcome was assumed.

Table A-3.3

Raw data of subject no.3. The participant was 36 years old, no student, and male. Duration of the experiment was 39 minutes.

chronology	video no.1	video no.2	video no.3	video no.4	video no.5	video no.6	video no.7	video no.8	video no.9
1st	.200	-.153	.184	-.145	-.129	-.184	-.373	-.145	.122
2nd	.357	.357	-.169	.200	.310	-.388	-.490	-.365	-.529
3rd	.161	.412	-.294	-.161	.176	.231	-.310	.043	-.278
4th	.106	.341	.451	.114	-.278	.271	.247	-.443	-.341
5th	.278	.153	.098	.278	.325	-.059	-.475	-.200	.404
6th	.114	-.278	-.106	.263	.137	.090	-.122	-.404	.239
7th	-.373	.600	.388	-.231	.145	-.271	-.137	-.412	-.333
8th	.208	-.098	.459	-.294	.231	.184	-.192	-.153	-.341
9th	.404	.459	.396	-.051	-.584	.302	-.318	-.325	.271
10th	-.349	-.608	.231	.294	-.114	.271	.200	.349	-.569

Note. The videos were played in random order, therefore the chronological order is an inner-item value only, and no impact to the total outcome was assumed.

Table A-3.4

Raw data of subject no.4. The participant was 23 years old, student, and female. Duration of the experiment was 47 minutes.

chronology	video no.1	video no.2	video no.3	video no.4	video no.5	video no.6	video no.7	video no.8	video no.9
1st	.396	.000	-.851	.004	-.898	.529	-.435	-.169	.725
2nd	.514	-.231	.827	.224	-.412	-.231	.420	-.427	.388
3rd	.427	.294	-.380	-.498	-.498	.859	.333	.000	.976
4th	.733	-.475	-.365	.200	.000	-.325	-.616	-.278	.984
5th	.200	.200	.929	-.216	-.875	.137	-.278	-.318	.451
6th	.773	.608	.404	.200	-.718	-.663	-.694	-.710	.357
7th	.545	.145	.561	-.945	-.537	-.435	.192	-.114	-.357
8th	.875	.271	-.820	.000	-.145	-.161	.184	-.859	-.247
9th	.875	-.357	-.271	.000	-.271	.231	.000	-.976	-.216
10th	.333	.216	-.247	.000	-.765	.451	.000	-.325	-.349

Note. The videos were played in random order, therefore the chronological order is an inner-item value only, and no impact to the total outcome was assumed.

Appendix A4 [Pre-Experiment I – Tables]

The tables in appendix A contain all tables used in the paper as well as additional tables referred to in the paper in the pre-experiment I section. All tables are digitally available in the MS Excel 97/2003 format [*.xls] and the LibreOffice Calc (version 4.0) format [*.ods] on the appendix CD [/content/files/appendix/01_pre-experiment1/tables/..].

Table A-4.1

Means (Mean Differences), Standard Deviations, Standard Error of Means, and statistics of the Mann-Whitney U Test (Mean Rank and Sum of Ranks).

	Group (student)	N[a]	M (MD)[b]	SD	SEM	Mean Rank[c]	Sum of Ranks[c]
video no.1	yes	20	.269	.360	.080	21.18	423.50
	no	20	.197	.279	.062	19.83	396.50
	Total	40	.233	.320	.051		
video no.2	yes	20	.030	.259	.058	17.00	340.00
	no	20	.182	.334	.075	24.00	480.00
	Total	40	.106	.305	.048		
video no.3	yes	20	-.034	.464	.104	16.73	334.50
	no	20	.188	.288	.064	24.28	485.50
	Total	40	.077	.398	.063		
video no.4	yes	20	-.019	.301	.067	19.88	397.50
	no	20	.032	.242	.054	21.13	422.50
	Total	40	.006	.270	.043		
video no.5	yes	20	-.245	.360	.081	18.28	365.50
	no	20	-.123	.377	.084	22.73	454.50
	Total	40	-.184	.369	.058		

(Table A-4.1 continues)

(Table A-4.1 continued)

	Group (student)	N[a]	M (MD)[b]	SD	SEM	Mean Rank[c]	Sum of Ranks[c]
video no.6	yes	20	.013	.356	.080	18.83	376.50
	no	20	.053	.299	.067	22.18	443.50
	Total	40	.033	.326	.051		
video no.7	yes	20	-.074	.289	.065	24.78	495.50
	no	20	-.255	.224	.050	16.23	324.50
	Total	40	-.164	.271	.043		
video no.8	yes	20	-.270	.286	.064	19.85	397.00
	no	20	-.159	.283	.063	21.15	423.00
	Total	40	-.215	.286	.045		
video no.9	yes	20	.144	.411	.092	22.50	450.00
	no	20	-.027	.406	.091	18.50	370.00
	Total	40	.058	.412	.065		

Note. The data is based on a continuous scale ranging from value -1.00 (defined as male/masculine) to 1.00 (defined as female/feminine).

a. The total number of subjects was 4 (2 students, 2 nonstudents); the sample size is related to the number of data sets (each video was shown 10 times to each subject).

b. Mean (M) and mean difference (MD) is the same for the (in this experiment) tested value $\mu_0 = .00$.

c. Statistics of the Mann-Whitney U Test.

Table A-4.2

Tests of Normality of the full sample for all 9 videos.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
video no.1	.057	40	.200*	.981	40	.726
video no.2	.105	40	.200*	.968	40	.304
video no.3	.102	40	.200*	.972	40	.417
video no.4	.109	40	.200*	.924	40	.010**
video no.5	.120	40	.147	.938	40	.031
video no.6	.107	40	.200*	.979	40	.638
video no.7	.063	40	.200*	.988	40	.943
video no.8	.136	40	.059	.950	40	.079
video no.9	.118	40	.175	.961	40	.178

Note. Due to the small sample size (N < 50) the Shapiro-Wilk Test was performed in addition to the Kolmogorov-Smirnov Test.

a. Lilliefors Significance Correction.

*This is a lower bound of the true significance, **p < .05

Table A-4.3

Mann-Whitney U Test comparing students and non-students.

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)	Exact Sig. [2*(1-tailed Sig.)]
video no.1	186.5	397	-.365	.715	.718[a]
video no.2	130.0	340	-1.894	.058	.060[a]
video no.3	124.5	335	-2.043	.041	.040[a]
video no.4	187.5	398	-.338	.735	.738[a]
video no.5	155.5	366	-1.205	.228	.231[a]
video no.6	166.5	377	-.906	.365	.369[a]
video no.7	114.5	325	-2.314	.021	.020[a]
video no.8	187.0	397	-.352	.725	.738[a]
video no.9	160.0	370	-1.082	.279	.289[a]

[a]Not corrected for ties.

Table A-4.4

Levene's Test for Equality of Variances comparing the sub-samples 'studens' and 'non-students' for all 9 videos.

	F	Sig.
video no.1	3.099	.086
video no.2	1.802	.187
video no.3	3.174	.083
video no.4	.160	.691
video no.5	.094	.760
video no.6	.544	.465
video no.7	.993	.325
video no.8	.016	.901
video no.9	.392	.535

Table A-4.5

Independent two-tailed t-test for equality of means of students and non-students.

	t	df	Sig. (2-tailed)	MD	SED	95% Confidence Interval of the Difference	
						Lower	Upper
video no.1	.710	38	.482	.072	.102	-.134	.279
video no.2	-1.608	38	.116	-.152	.094	-.343	.039
video no.3	-1.816	38	.077	-.222	.122	-.469	.025
video no.4	-.589	38	.559	-.051	.086	-.225	.124
video no.5	-1.045	38	.303	-.122	.117	-.358	.114
video no.6	-.380	38	.706	-.040	.104	-.250	.171
video no.7	2.206*	38	.034	.180	.082	.015	.346
video no.8	-1.235	38	.224	-.111	.090	-.293	.071
video no.9	1.317	38	.196	.170	.129	-.091	.432

*$p < .05$

Table A-4.6

Two-tailed one-sample t-test for $\bar{x}_{tot}$ (full sample).

	Test Value = .00					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	M (MD)	Lower	Upper
video no.1	4.610***	39	.000	.233	.131	.336
video no.2	2.199*	39	.034	.106	.009	.203
video no.3	1.222	39	.229	.077	-.050	.204
video no.4	.149	39	.882	.006	-.080	.093
video no.5	-3.147**	39	.003	-.184	-.302	-.066
video no.6	.636	39	.528	.033	-.071	.137
video no.7	-3.833***	39	.000	-.164	-.251	-.078
video no.8	-4.750***	39	.000	-.215	-.306	-.123
video no.9	.896	39	.376	.058	-.073	.190

Note. N=40 (10 times each video per subject). The data is based on a continuous scale ranging from -1.0 (male/masculine) to 1.0 (female/feminine). Mean (M) and mean difference (MD) is the same for the test value $\mu_0 = .00$.

*p < .05, **p < .01, ***p < .001

Table A-4.7

Two-tailed one-sample t-test for $\bar{x}_{stud}$ (students).

	Test Value = .00					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	M (MD)	Lower	Upper
video no.1	3.350*	19	.003	.269	.101	.438
video no.2	.519	19	.610	.030	-.091	.151
video no.3	-.329	19	.746	-.034	-.251	.183
video no.4	-.283	19	.780	-.019	-.160	.122
video no.5	-3.036*	19	.007	-.245	-.413	-.076
video no.6	.162	19	.873	.013	-.154	.180
video no.7	-1.147	19	.266	-.074	-.209	.061
video no.8	-4.234**	19	.000	-.270	-.404	-.137
video no.9	1.560	19	.135	.144	-.049	.336

Note. N=20 (10 times each video per subject). The data is based on a continuous scale ranging from -1.0 (male/masculine) to 1.0 (female/feminine). Mean (M) and mean difference (MD) is the same for the test value $\mu_0 = .00$.

*p < .01, **p < .001

Table A-4.8

Two-tailed one-sample t-test for $\bar{x}_{\neg stud}$ (non-students).

	Test Value = .00					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	M (MD)	Lower	Upper
video no.1	3.154**	19	.005	.197	.066	.328
video no.2	2.435*	19	.025	.182	.026	.338
video no.3	2.914**	19	.009	.188	.053	.323
video no.4	.588	19	.564	.032	-.081	.145
video no.5	-1.456	19	.162	-.123	-.299	.054
video no.6	.785	19	.442	.053	-.088	.193
video no.7	-5.076***	19	.000	-.255	-.359	-.150
video no.8	-2.521*	19	.021	-.159	-.292	-.027
video no.9	-.294	19	.772	-.027	-.217	.163

Note. N=20 (10 times each video per subject). The data is based on a continuous scale ranging from -1.0 (male/masculine) to 1.0 (female/feminine). Mean (M) and mean difference (MD) is the same for the test value $\mu_0 = .00$.

*p < .05, **p < .01, ***p < .001

Appendix A5 [Pre-Experiment I – Figures]

The figures in appendix A contain all figures used in-the paper as well as additional figures referred to in the paper in the pre-experiment I section. All figures are digitally available in the Portable Network Graphic format [*.png] on the appendix CD [/content/files/appendix/01_pre-experiment1/figures/..].

Box Plots

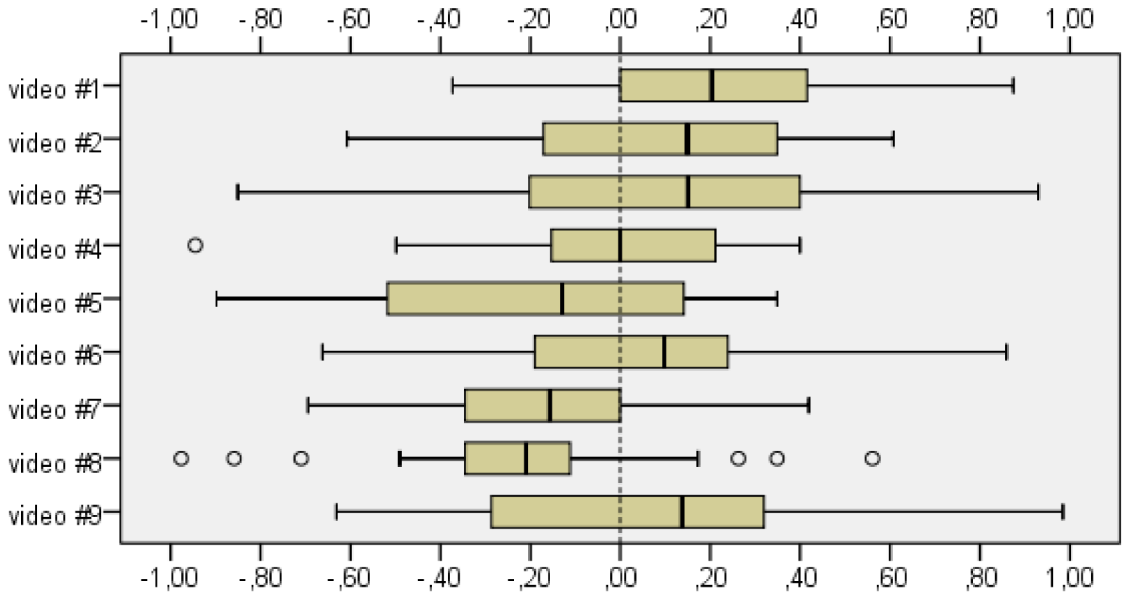


Figure A-5.1. Box Plot for full sample (N = 40) gender ratings for all 9 videos.

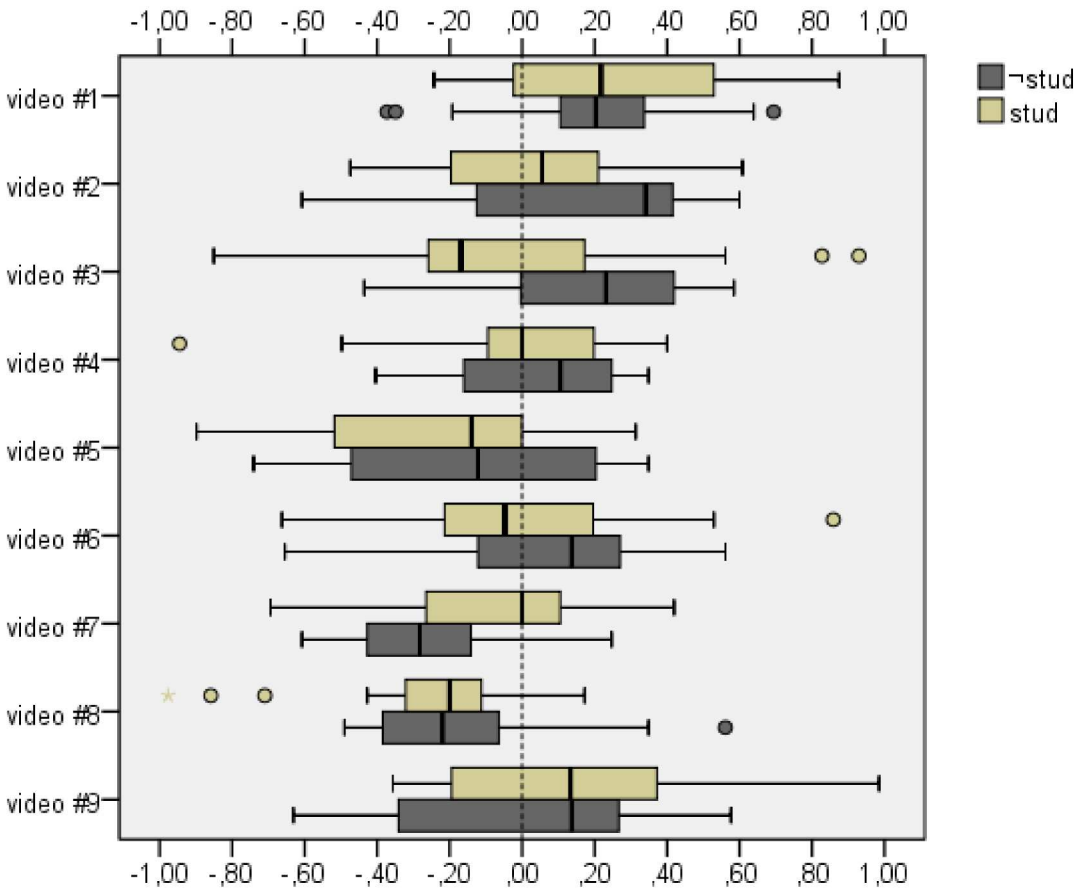


Figure A-5.2. Box Plot comparing the two sub-samples, students (N = 20) and non-students (N = 20), to each other for all 9 videos.

Histograms and Q-Q Plots

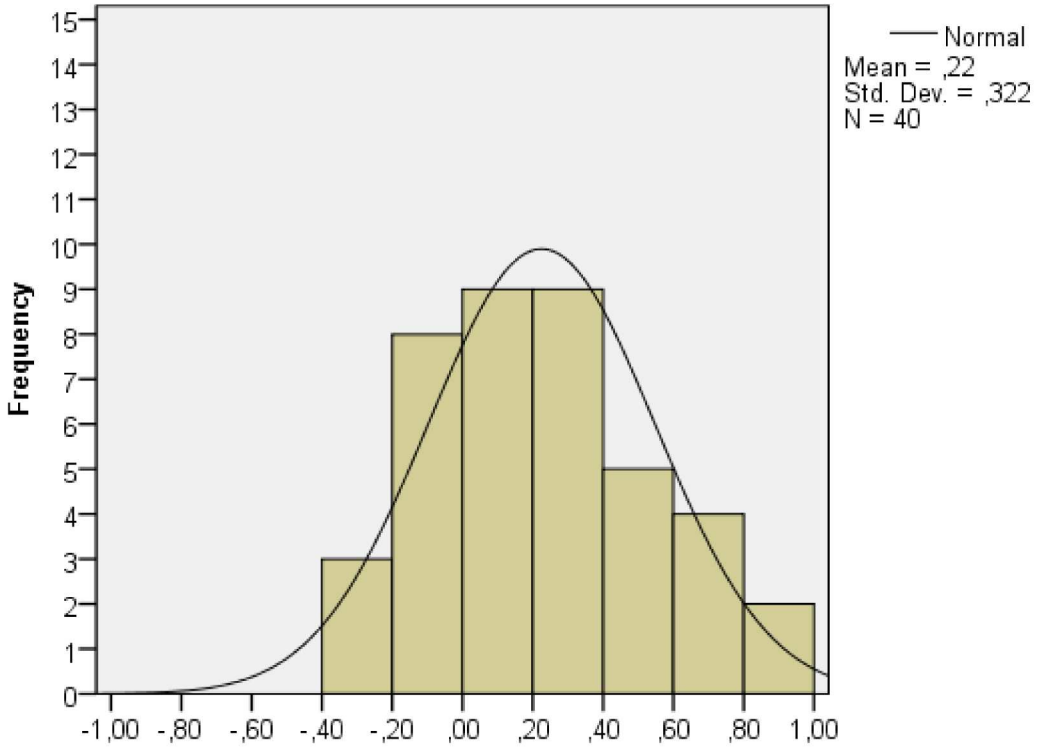


Figure A-5.3. Histogram of gender ratings by the full sample (N = 40) for video no.1, classified in steps of 0.2.

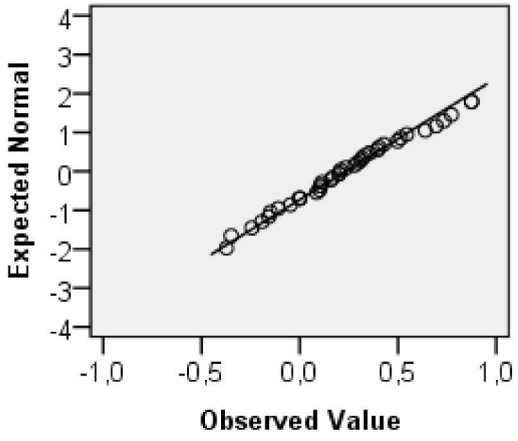


Figure A-5.3a. Normal Q-Q Plot of gender ratings for video no.1.

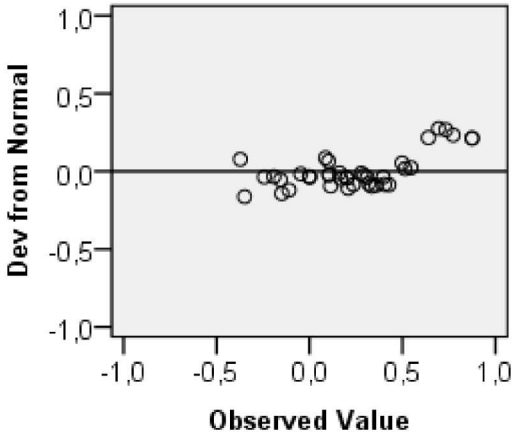


Figure A-5.3b. Detrended Normal Q-Q Plot of gender ratings for video no.1.

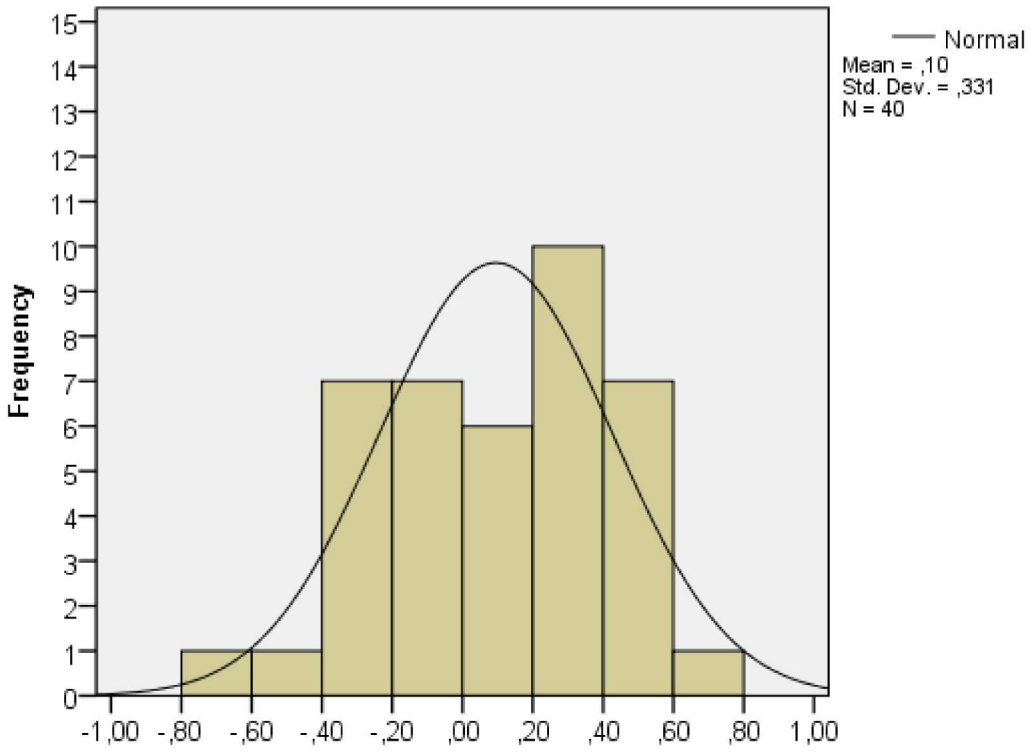


Figure A-5.4. Histogram of gender ratings by the full sample (N = 40) for video no.2, classified in steps of 0.2.

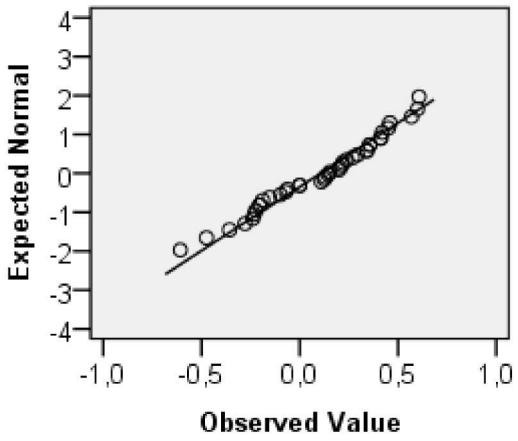


Figure A-5.4a. Normal Q-Q Plot of gender ratings for video no.2.

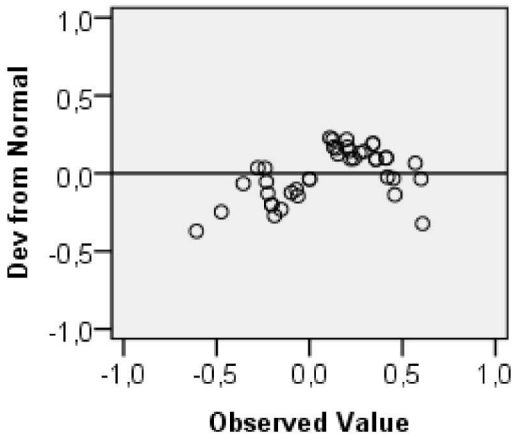


Figure A-5.4b. Detrended Normal Q-Q Plot of gender ratings for video no.2.

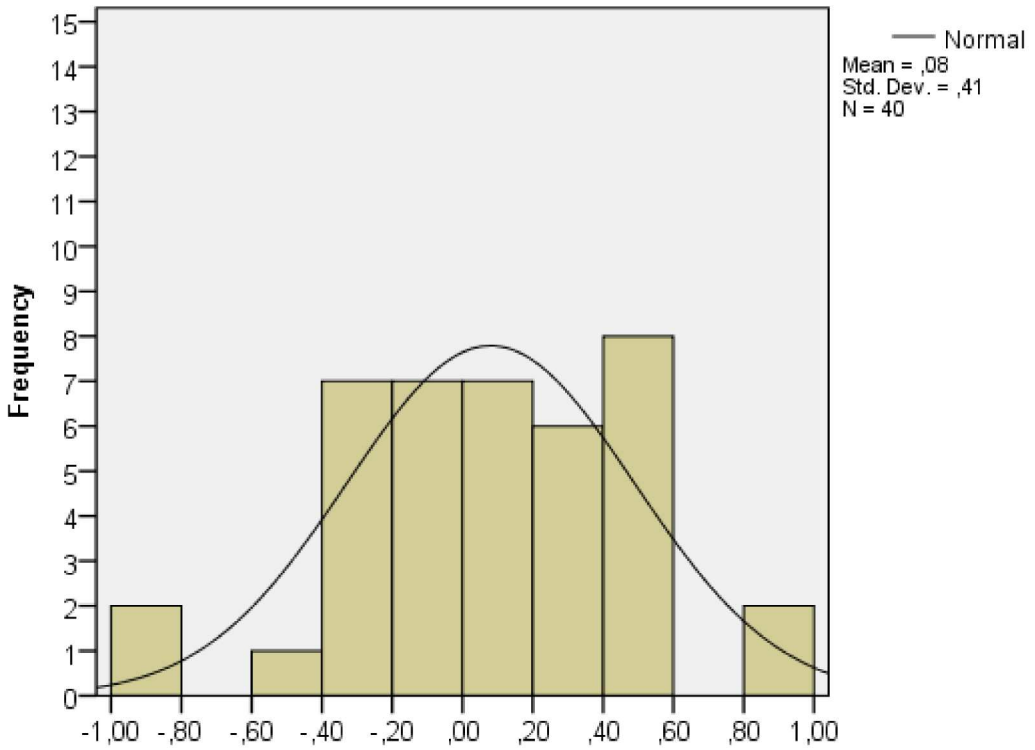


Figure A-5.5. Histogram of gender ratings by the full sample (N = 40) for video no.3, classified in steps of 0.2.

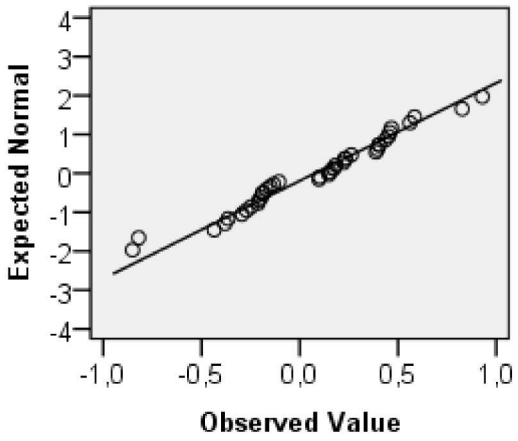


Figure A-5.5a. Normal Q-Q Plot of gender ratings for video no.3.

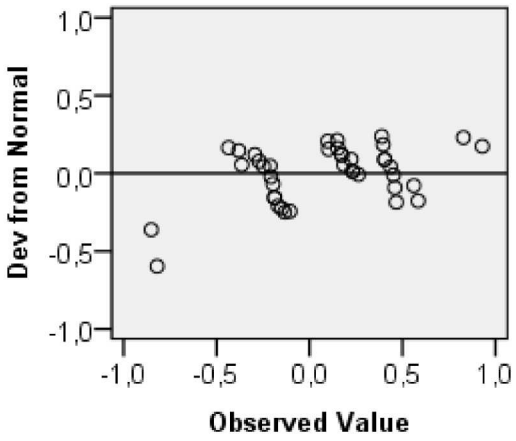


Figure A-5.5b. Detrended Normal Q-Q Plot of gender ratings for video no.3.

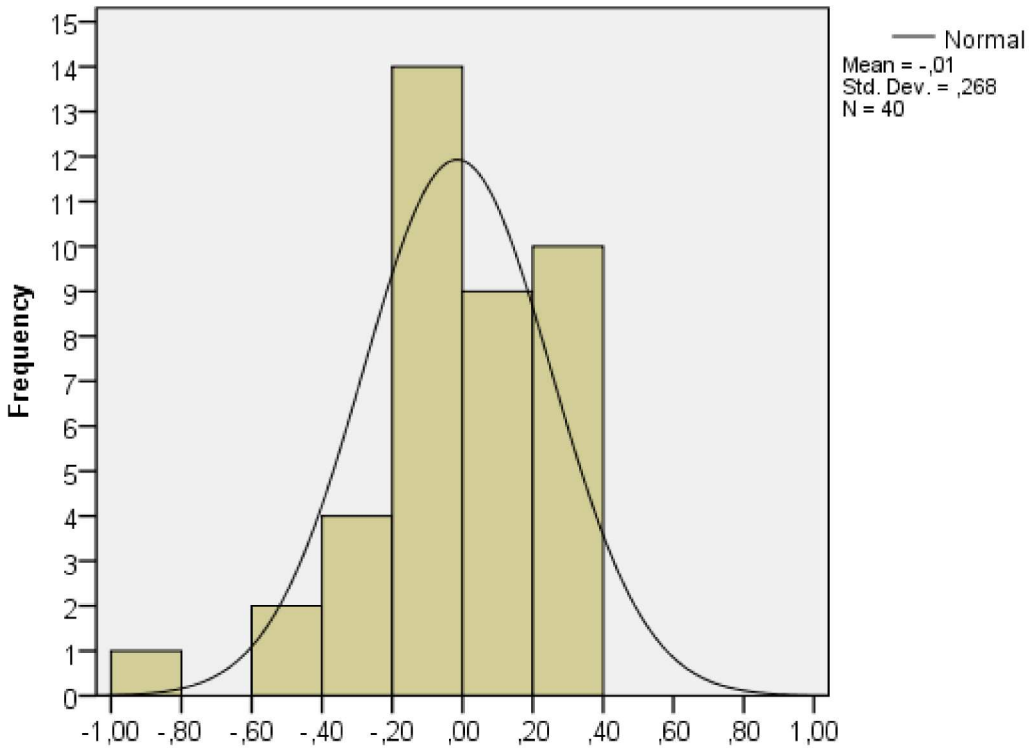


Figure A-5.6. Histogram of gender ratings by the full sample (N = 40) for video no.4, classified in steps of 0.2.

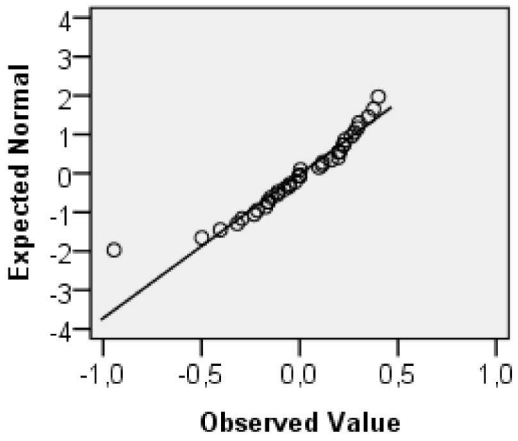


Figure A-5.6a. Normal Q-Q Plot of gender ratings for video no.4.

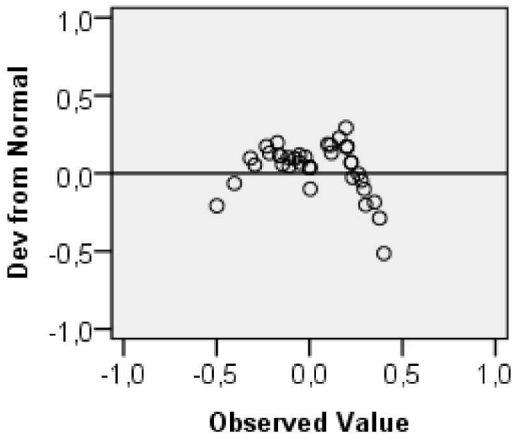


Figure A-5.6b. Detrended Normal Q-Q Plot of gender ratings for video no.4.

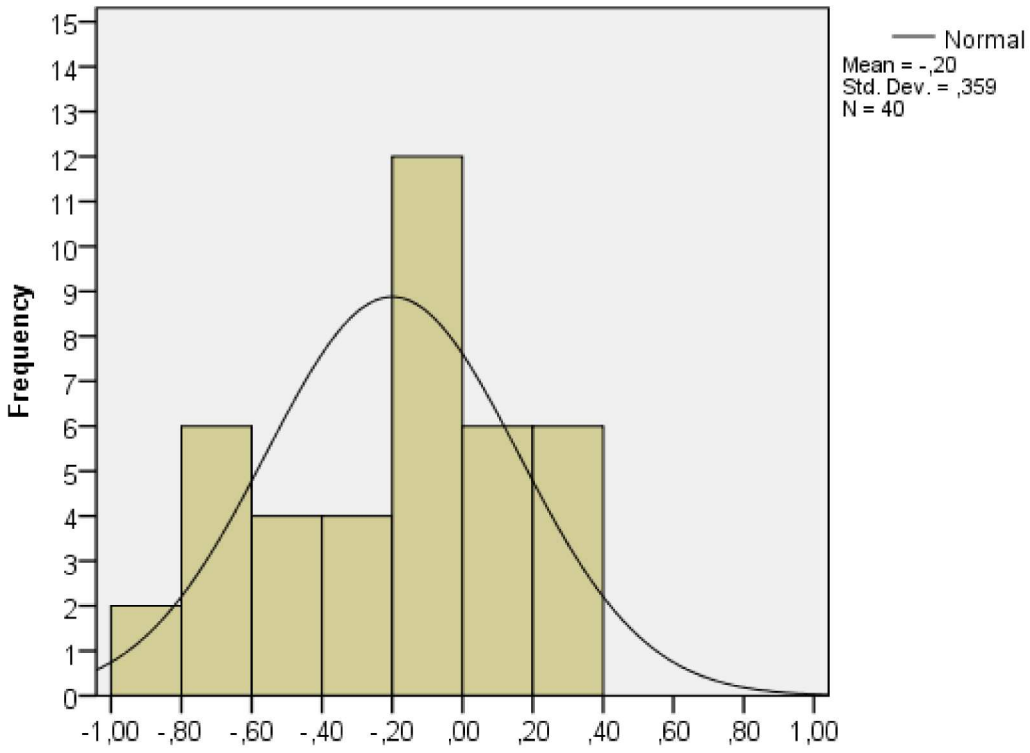


Figure A-5.7. Histogram of gender ratings by the full sample (N = 40) for video no.5, classified in steps of 0.2.

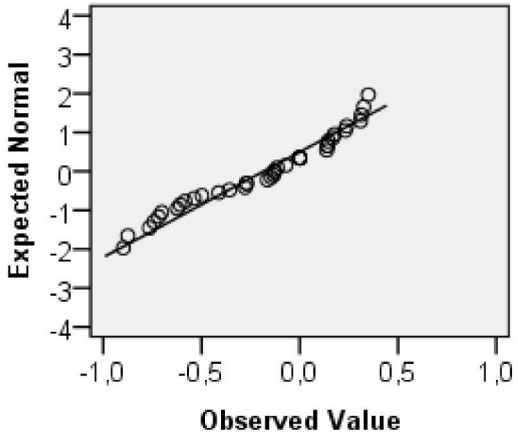


Figure A-5.7a. Normal Q-Q Plot of gender ratings for video no.5.

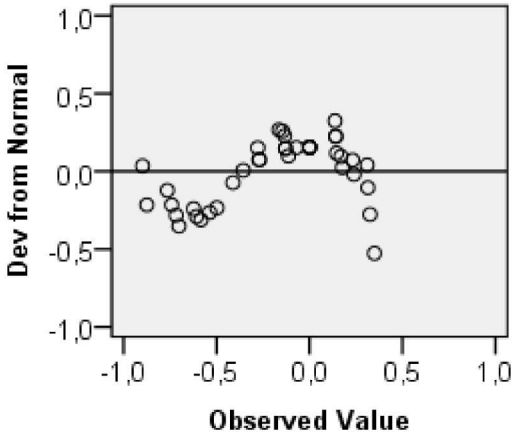


Figure A-5.7b. Detrended Normal Q-Q Plot of gender ratings for video no.5.

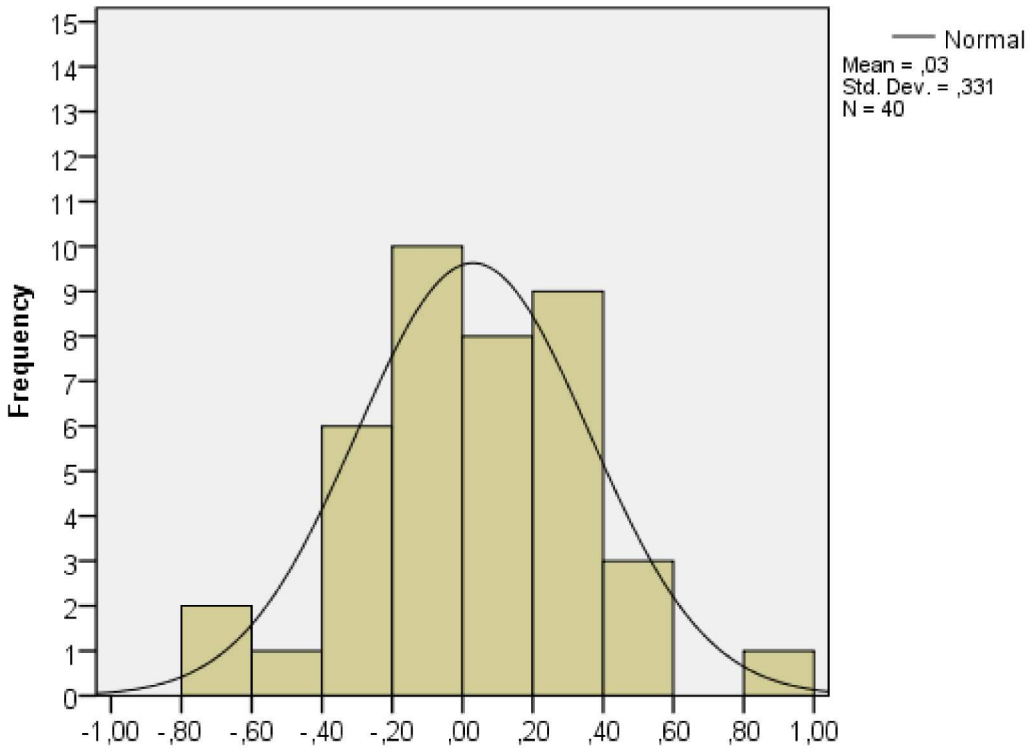


Figure A-5.8. Histogram of gender ratings by the full sample (N = 40) for video no.6, classified in steps of 0.2.

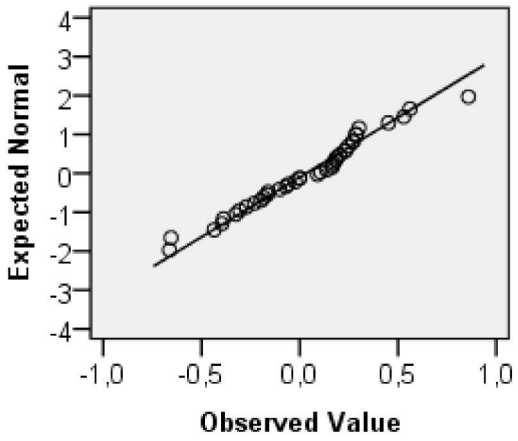


Figure A-5.8a. Normal Q-Q Plot of gender ratings for video no.6.

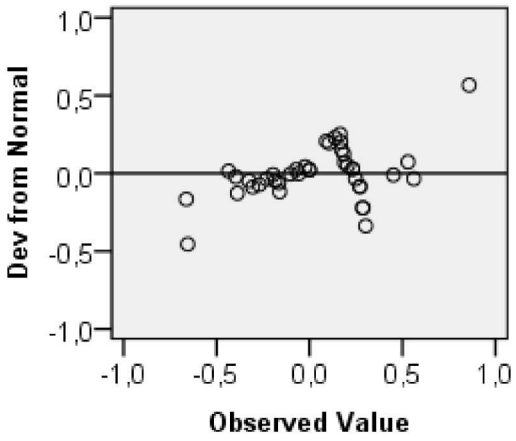


Figure A-5.8b. Detrended Normal Q-Q Plot of gender ratings for video no.6.

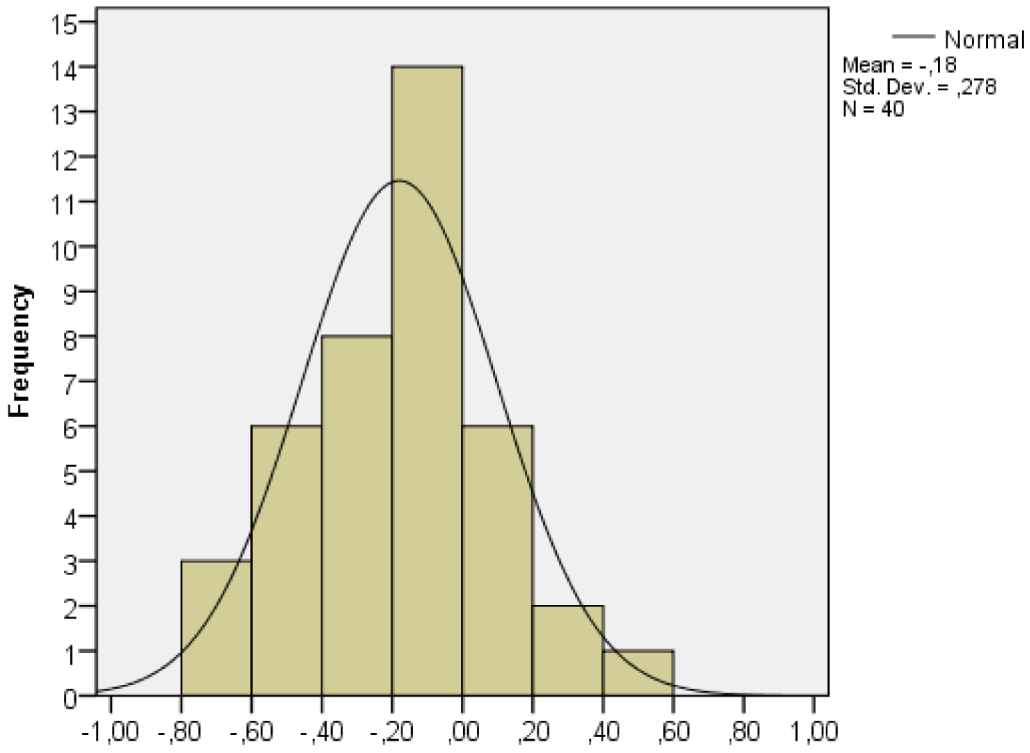


Figure A-5.9. Histogram of gender ratings by the full sample (N = 40) for video no.7, classified in steps of 0.2.

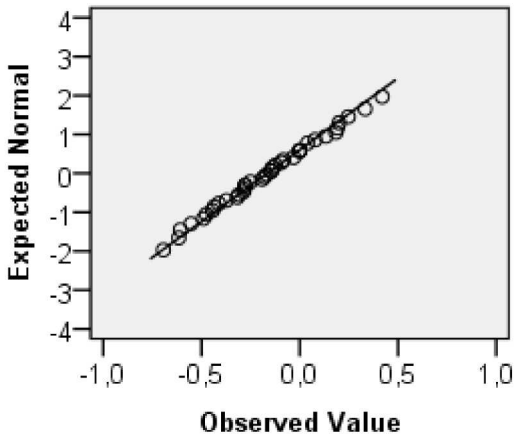


Figure A-5.9a. Normal Q-Q Plot of gender ratings for video no.7.

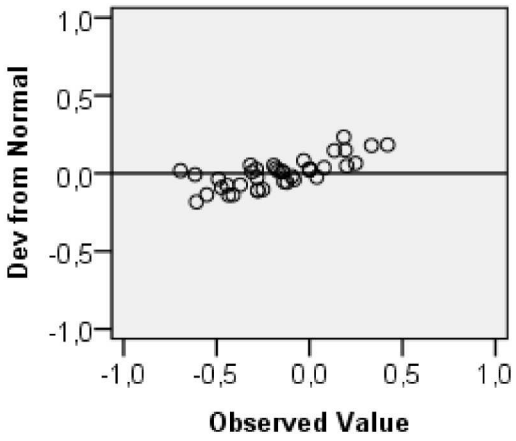


Figure A-5.9b. Detrended Normal Q-Q Plot of gender ratings for video no.7.

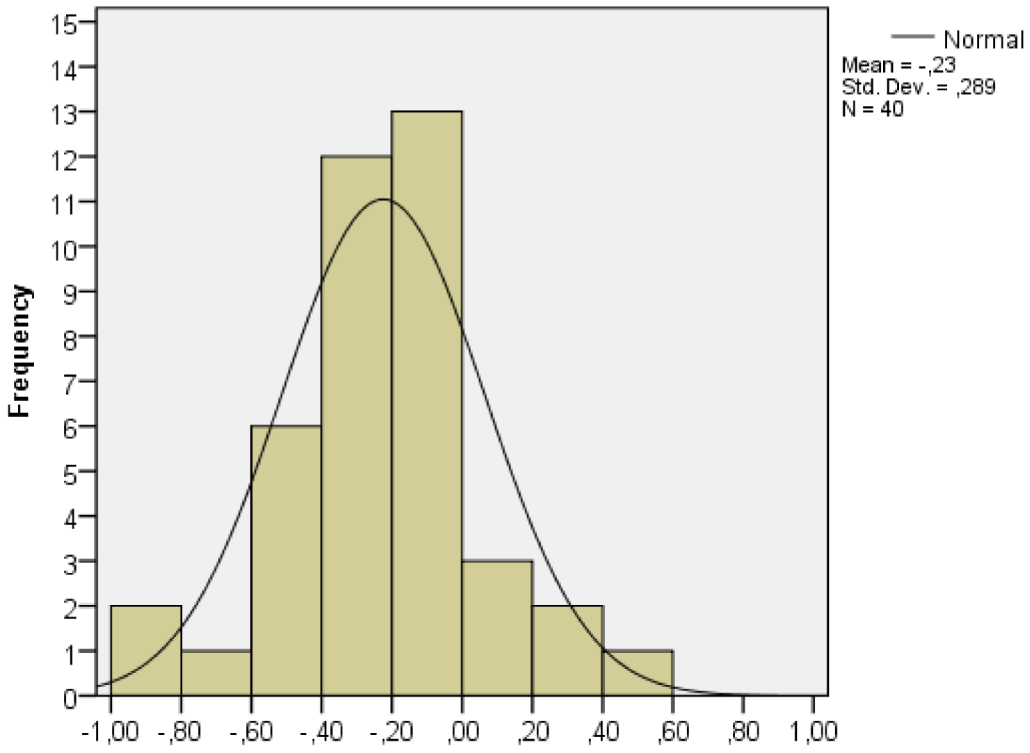


Figure A-5.10. Histogram of gender ratings by the full sample (N = 40) for video no.8, classified in steps of 0.2.

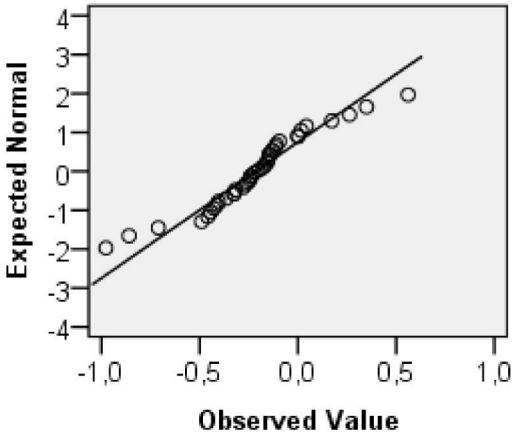


Figure A-5.10a. Normal Q-Q Plot of gender ratings for video no.9.

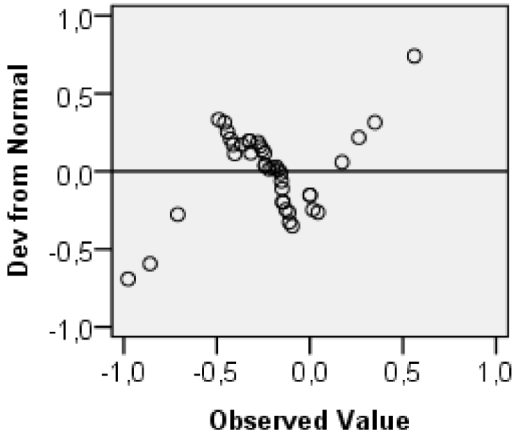


Figure A-5.10b. Detrended Normal Q-Q Plot of gender ratings for video no.9.

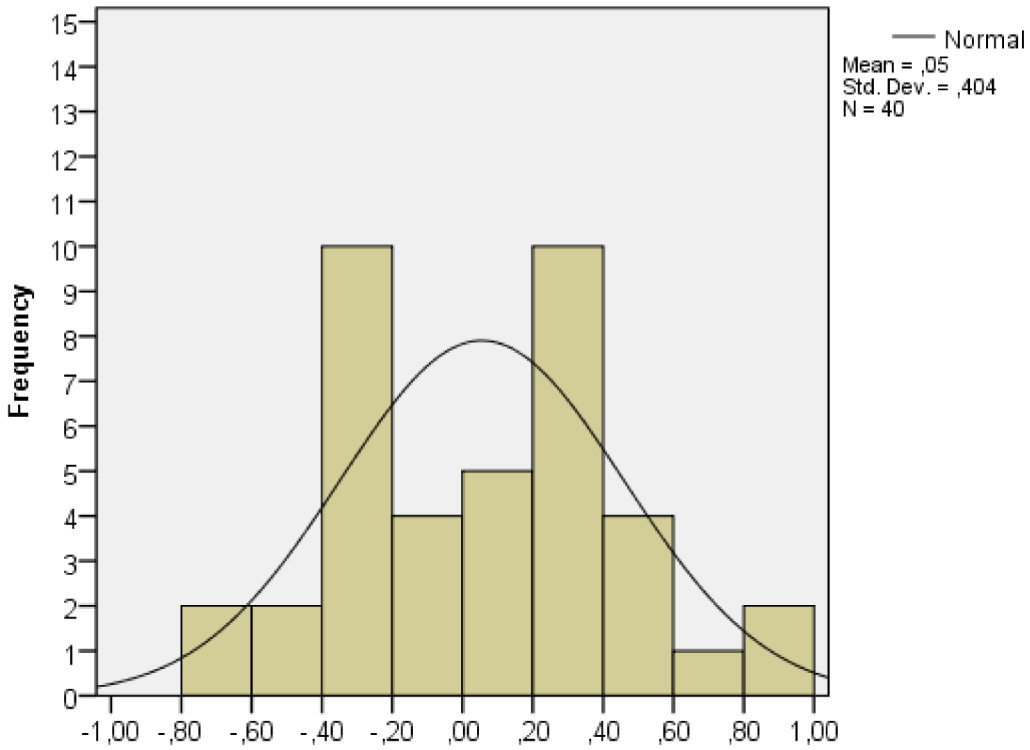


Figure A-5.11. Histogram of gender ratings by the full sample (N = 40) for video no.9, classified in steps of 0.2.

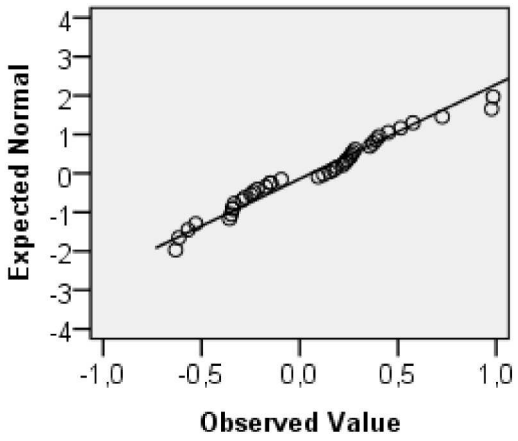


Figure A-5.11a. Normal Q-Q Plot of gender ratings for video no.9.

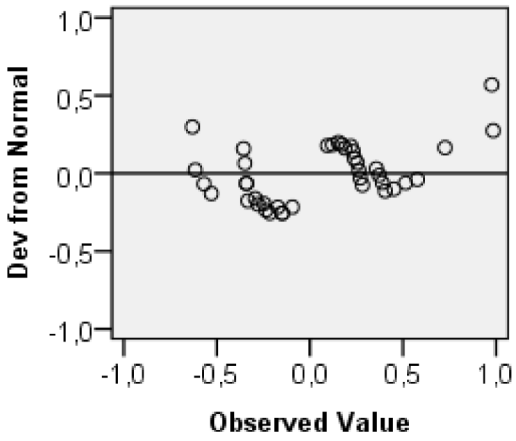
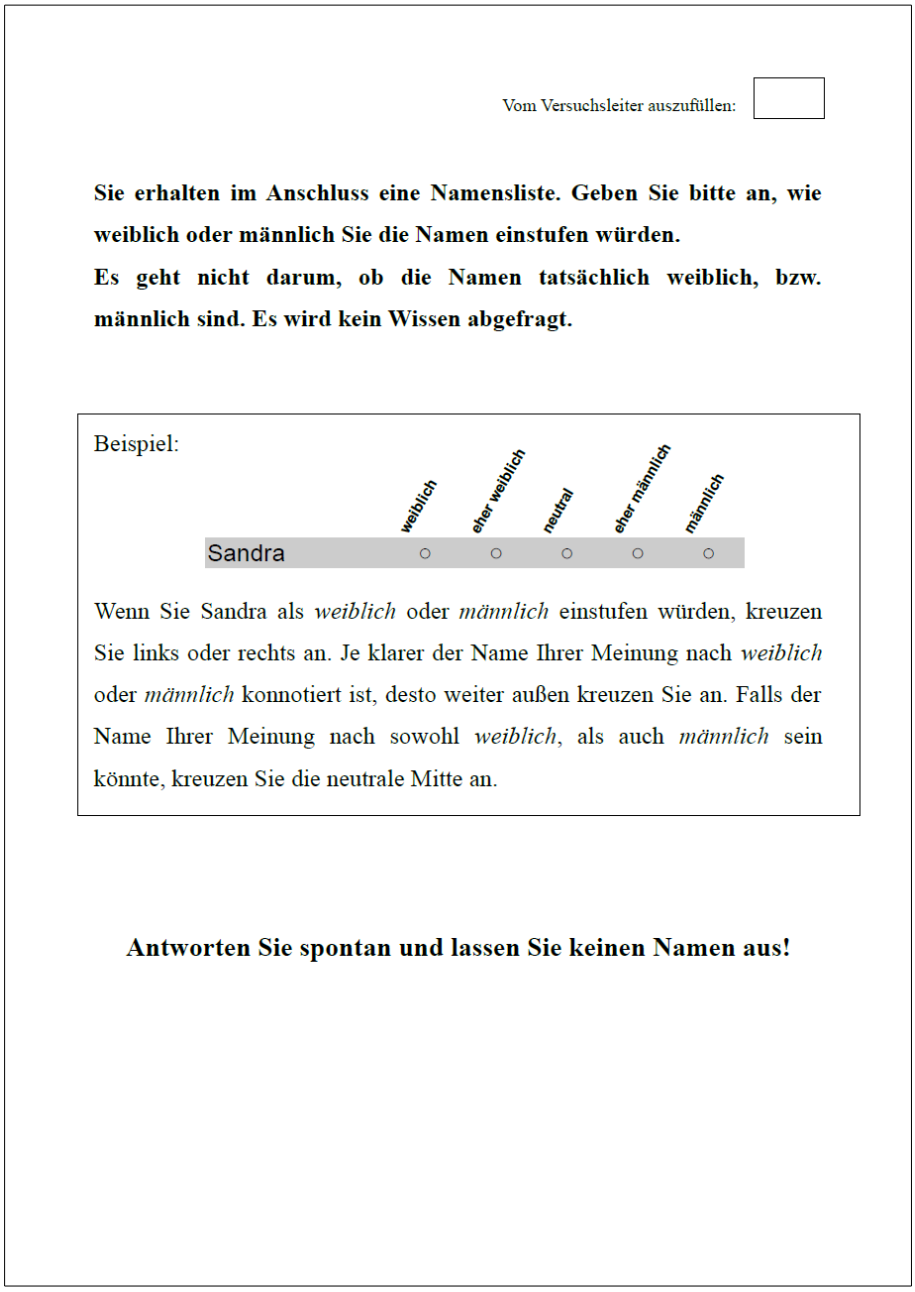


Figure A-5.11b. Detrended Normal Q-Q Plot of gender ratings for video no.9.

Appendix B0 [Pre-Experiment II – Instructions]



Vom Versuchsleiter auszufüllen:

Sie erhalten im Anschluss eine Namensliste. Geben Sie bitte an, wie weiblich oder männlich Sie die Namen einstufen würden.
Es geht nicht darum, ob die Namen tatsächlich weiblich, bzw. männlich sind. Es wird kein Wissen abgefragt.

Beispiel:

	weiblich	eher weiblich	neutral	eher männlich	männlich
Sandra	○	○	○	○	○

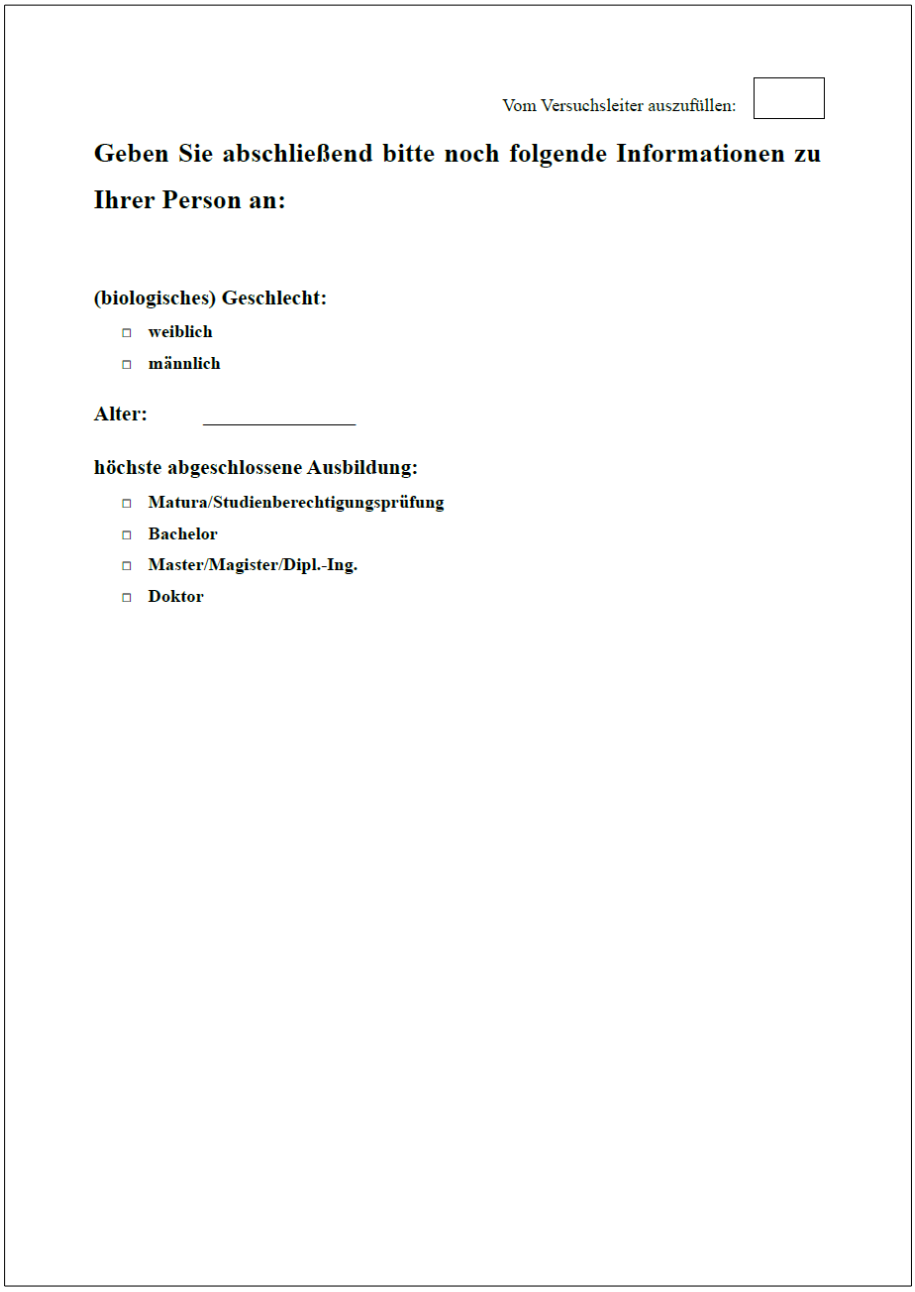
Wenn Sie Sandra als *weiblich* oder *männlich* einstufen würden, kreuzen Sie links oder rechts an. Je klarer der Name Ihrer Meinung nach *weiblich* oder *männlich* konnotiert ist, desto weiter außen kreuzen Sie an. Falls der Name Ihrer Meinung nach sowohl *weiblich*, als auch *männlich* sein könnte, kreuzen Sie die neutrale Mitte an.

Antworten Sie spontan und lassen Sie keinen Namen aus!

Figure B-0.1. Screenshot of the Adobe PDF file of the experiment instructions.

Original Adobe PDF available on appendix CD:
[/content/experiments/pre-experiment_#2/instructions.pdf]

Appendix B1 [Pre-Experiment II – Demographic Questionnaire]



Vom Versuchsleiter auszufüllen: ☐

Geben Sie abschließend bitte noch folgende Informationen zu Ihrer Person an:

(biologisches) Geschlecht:

- ☐ **weiblich**
- ☐ **männlich**

Alter: _____

höchste abgeschlossene Ausbildung:

- ☐ **Matura/Studienberechtigungsprüfung**
- ☐ **Bachelor**
- ☐ **Master/Magister/Dipl.-Ing.**
- ☐ **Doktor**

Figure B-1.1. Screenshot of the Adobe PDF file of the demographic questionnaire.

Original Adobe PDF available on appendix CD:
[/content/experiments/pre-experiment_#2/demographic.pdf]

Appendix B2 [Pre-Experiment II – Raw Data]

This appendix only includes the range, minimum and maximum rating, as well as the mean for each name. The raw data is included and available in SPSS file format [*.sav], and Adobe PDF[75] format [*.pdf] on the appendix CD [/content/experiments/pre-experiment_#2/data/..].

Table B-2.1

Overall view of the collected data (mean, range, minimum, and maximum ratings).

	N	Range	Minimum	Maximum	M
Aaron	10	1	1	2	1.90
Abigail	10	1	-2	-1	-1.70
Adam	10	1	1	2	1.90
Alani	10	4	-2	2	-.50
Alexander	10	1	1	2	1.80
Alexandra	10	3	-2	1	-1.50
Alina	10	2	-2	0	-1.80
Amelie	10	1	-2	-1	-1.90
Andreas	10	3	-1	2	1.60
Anna	10	3	-2	1	-1.70
Arif	10	2	0	2	1.20
Arthur	10	1	1	2	1.90
Astrid	10	4	-2	2	-1.40
Avery	10	4	-2	2	.00
Ayse	10	4	-2	2	-.60
Ben	10	2	0	2	1.70
Benjamin	10	1	1	2	1.90

(Table B-2.1 continues)

75 The newest version of Adobe Acrobat Reader® can be downloaded for free at http://get.adobe.com/de/reader/

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Bianca	10	1	-2	-1	-1.90
Bircan	10	2	0	2	1.30
Cameron	10	4	-2	2	.60
Cengiz	10	2	0	2	1.40
Christine	10	1	-2	-1	-1.70
Dace	10	3	-1	2	.40
Daniel	10	1	1	2	1.90
David	10	2	0	2	1.80
Denise	10	4	-2	2	-1.30
Deniz	10	4	-2	2	1.20
Dennis	10	2	0	2	1.70
Dina	10	2	-2	0	-1.60
Diniz	10	3	-2	1	-.40
Dominique	10	2	-2	0	-1.00
Dylan	10	3	-1	2	1.30
Edvin	10	1	1	2	1.80
Elena	10	1	-2	-1	-1.90
Elias	10	3	-1	2	1.60
Elise	10	4	-2	2	-1.40
Elvira	10	1	-2	-1	-1.70
Emilia	10	0	-2	-2	-2.00
Emma	10	3	-2	1	-1.50
Erik	10	0	2	2	2.00
Erin	10	4	-2	2	-.30
Esau	10	3	-1	2	1.00
Ethan	10	2	0	2	1.70

(Table B-2.1 continues)

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Eva	10	3	-2	1	-1.70
Evan	10	2	0	2	1.40
Ezra	10	4	-2	2	-.40
Fabian	10	1	1	2	1.70
Faizah	10	3	-2	1	-.70
Felix	10	1	1	2	1.90
Finlay	10	3	-1	2	1.20
Finn	10	1	1	2	1.70
Fiona	10	4	-2	2	-1.50
Florence	10	2	-2	0	-1.60
Florian	10	2	0	2	1.80
Franz	10	0	2	2	2.00
Freja	10	2	-2	0	-1.20
Gan	10	2	0	2	.80
Gerhard	10	0	2	2	2.00
Gideon	10	2	0	2	1.60
Gil	10	2	0	2	.50
Grace	10	1	-2	-1	-1.70
Guilherme	10	4	-2	2	.30
Gustav	10	0	2	2	2.00
Haakon	10	1	1	2	1.70
Hamid	10	1	1	2	1.70
Hannah	10	1	-2	-1	-1.80
Hoa	10	3	-1	2	.40
Hugo	10	0	2	2	2.00
Hung	10	2	0	2	.70

(Table B-2.1 continues)

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Ida	10	1	-2	-1	-1.70
Ingrid	10	1	-2	-1	-1.80
Irfan	10	4	-2	2	.50
Isabell	10	0	-2	-2	-2.00
Jade	10	4	-2	2	-1.00
Jaime	10	4	-2	2	.40
Jan	10	1	1	2	1.90
Jana	10	0	-2	-2	-2.00
Janis	10	4	-2	2	-.10
Jemina	10	2	-2	0	-1.50
Johannes	10	0	2	2	2.00
Jonas	10	1	1	2	1.80
José	10	2	0	2	1.60
Josua	10	1	1	2	1.80
Julia	10	0	-2	-2	-2.00
Julian	10	3	-1	2	1.60
Justin	10	1	1	2	1.90
Katharina	10	1	-2	-1	-1.90
Keiki	10	4	-2	2	-.20
Kim	10	4	-2	2	-.50
Klemens	10	1	1	2	1.90
Kyle	10	2	0	2	1.30
Lana	10	1	-2	-1	-1.90
Lara	10	0	-2	-2	-2.00
Larissa	10	0	-2	-2	-2.00
Lærke	10	3	-1	2	.40

(Table B-2.1 continues)

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Latif	10	3	-1	2	.50
Laura	10	1	-2	-1	-1.80
Lena	10	1	-2	-1	-1.90
Leon	10	2	0	2	1.80
Leonie	10	1	-2	-1	-1.90
Leyla	10	0	-2	-2	-2.00
Liam	10	3	-1	2	1.20
Luoanne	10	1	-2	-1	-1.60
Louis	10	4	-2	2	1.20
Lovro	10	2	0	2	1.00
Lucie	10	1	-2	-1	-1.80
Lucy	10	1	-2	-1	-1.90
Lukas	10	0	2	2	2.00
Madison	10	4	-2	2	-.40
Makani	10	3	-1	2	-.10
Mara	10	2	-2	0	-1.70
Marcel	10	1	1	2	1.70
Marek	10	1	1	2	1.70
Marie	10	1	-2	-1	-1.80
Marit	10	4	-2	2	-1.00
Markus	10	0	2	2	2.00
Matej	10	2	0	2	1.60
Matthias	10	0	2	2	2.00
Maxime	10	4	-2	2	-.40
Maximilian	10	0	2	2	2.00
Maya	10	2	-2	0	-1.80

(Table B-2.1 continues)

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Mercedes	10	4	-2	2	-.60
Mia	10	1	-2	-1	-1.90
Michal	10	3	-1	2	1.40
Mikkel	10	2	0	2	1.80
Moritz	10	0	2	2	2.00
Nadim	10	4	-2	2	.60
Nadine	10	0	-2	-2	-2.00
Natalie	10	1	-2	-1	-1.90
Nathan	10	1	1	2	1.60
Nico	10	2	0	2	1.60
Nikolaj	10	1	1	2	1.90
Nina	10	0	-2	-2	-2.00
Noah	10	2	0	2	1.50
Oliver	10	1	1	2	1.90
Olivia	10	1	-2	-1	-1.80
Olle	10	2	0	2	1.30
Ondrej	10	2	0	2	1.50
Osman	10	0	2	2	2.00
Otto	10	1	1	2	1.90
Patrick	10	3	-1	2	1.70
Pedro	10	1	1	2	1.90
Pinar	10	4	-2	2	-.10
Pua	10	3	-1	2	-.10
Pualani	10	3	-1	2	.30
Rachel	10	2	-2	0	-1.60
Richard	10	0	2	2	2.00

(Table B-2.1 continues)

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Rita	10	1	-2	-1	-1.80
Robin	10	2	0	2	1.10
Ronja	10	3	-2	1	-1.50
Roque	10	2	0	2	1.50
Ryan	10	2	0	2	1.70
Sabri	10	3	-2	1	-.80
Safa	10	3	-2	1	-.50
Saga	10	3	-2	1	-.50
Sarah	10	1	-2	-1	-1.90
Sascha	10	2	0	2	1.10
Sebastian	10	3	-1	2	1.70
Seher	10	4	-2	2	.50
Selina	10	1	-2	-1	-1.80
Seval	10	2	0	2	1.00
Sidney	10	3	-2	1	-.40
Simon	10	1	1	2	1.90
Simone	10	3	-2	1	-1.30
Sofia	10	1	-2	-1	-1.90
Sophie	10	1	-2	-1	-1.90
Susanne	10	1	-2	-1	-1.90
Sven	10	0	2	2	2.00
Tahir	10	2	0	2	1.30
Tanja	10	3	-2	1	-1.70
Tarek	10	2	0	2	1.60
Thao	10	2	0	2	1.10
Tim	10	1	1	2	1.80

(Table B-2.1 continues)

(Table B-2.1 continued)

	N	Range	Minimum	Maximum	M
Tobias	10	1	1	2	1.90
Tom	10	1	1	2	1.90
Tung	10	2	0	2	.80
Tuva	10	3	-1	2	.00
Vega	10	4	-2	2	.30
Vera	10	1	-2	-1	-1.90
Vilde	10	3	-1	2	.00
Wedat	10	2	0	2	.80
William	10	3	-1	2	1.70
Yael	10	3	-1	2	.60
Yanis	10	4	-2	2	1.00
Yasar	10	2	0	2	1.10
Zada	10	2	-2	0	-.90

Appendix B3 [Pre-Experiment II – Tables]

The tables in appendix B contain all tables used in the paper as well as additional tables referred to in the paper in the pre-experiment II section. All tables are digitally available in the MS Excel 97/2003 format [*.xls] and the LibreOffice Calc (version 4.0) format [*.ods] on the appendix CD [/content/files/appendix/02_pre-experiment2/tables/..].

Table B-3.1

List of names used in the main study to label the motion capture videos in alphabetic order. The table contains the primary selection criteria (means, standard error, and standard deviations), as well as the range and the number of valid ratings of the second pre-experiment.

Name	N	Range	M	SE	SD
Aaron	10	1	1.90	.100	.316
Adam	10	1	1.90	.100	.316
Arthur	10	1	1.90	.100	.316
Benjamin	10	1	1.90	.100	.316
Bianca	10	1	-1.90	.100	.316
Daniel	10	1	1.90	.100	.316
Elena	10	1	-1.90	.100	.316
Emilia	10	0	-2.00	.000	.000
Erik	10	0	2.00	.000	.000
Gerhard	10	0	2.00	.000	.000
Isabell	10	0	-2.00	.000	.000

(Table B-3.1 continues)

(Table B-3.1 continued)

Name	N	Range	M	SE	SD
Jamie[a]	-	-	-	-	-
Jana	10	0	-2.00	.000	.000
Julia	10	0	-2.00	.000	.000
Katharina	10	1	-1.90	.100	.316
Lana	10	1	-1.90	.100	.316
Lara	10	0	-2.00	.000	.000
Larissa	10	0	-2.00	.000	.000
Lee[a]	-	-	-	-	-
Lena	10	1	-1.90	.100	.316
Leyla	10	0	-2.00	.000	.000
Lucy	10	1	-1.90	.100	.316
Lukas	10	0	2.00	.000	.000
Maemi-Haru[a]	-	-	-	-	-
Markus	10	0	2.00	.000	.000
Matthias	10	0	2.00	.000	.000
Maximilian	10	0	2.00	.000	.000
Mia	10	1	-1.90	.100	.316
Nadine	10	0	-2.00	.000	.000
Natalie	10	1	-1.90	.100	.316
Nikolaj	10	1	1.90	.100	.316

(Table B-3.1 continues)

(Table B-3.1 continued)

Name	N	Range	M	SE	SD
Nina	10	0	-2.00	.000	.000
Oliver	10	1	1.90	.100	.316
Osman	10	0	2.00	.000	.000
Pedro	10	1	1.90	.100	.316
Richard	10	0	2.00	.000	.000
Robin[b]	- (10)	- (2)	- (1.10)	- (.314)	- (.994)
Sam[a]	-	-	-	-	-
Sasha[a]	-	-	-	-	-
Sophie	10	1	-1.90	.100	.316
Summer[a]	-	-	-	-	-
Sven	10	0	2.00	.000	.000
Tom	10	1	1.90	.100	.316
Uli[a]	-	-	-	-	-
Ying-Yu[a]	-	-	-	-	-
Valid N (listwise)	10				

Note. Only male and female names used in the main study were tested.

a. Neutral names that were taken out of databases.

b. The only name taken randomly but also appeared in the test names.

Achievement Tables

Table B-3.2

Achievement levels and t-values for $x_p \equiv \mu_0 = -2.00$ (perfect female mean).

	A(x,y)	t	Sig. (1-tailed)	M	SD
Aaron	.047	39.000	.000	1.90	.316
Abigail	.848	1.964	.041	-1.70	.483
Adam	.047	39.000	.000	1.90	.316
Alani	.184	3.737	.002	-.50	1.269
Alexander	.048	28.500	.000	1.80	.422
Alexandra	.592	1.464	.089	-1.50	1.080
Alina	.869	1.000	.172	-1.80	.632
Amelie	.982	1.000	.172	-1.90	.316
Andreas	.046	11.784	.000	1.60	.966
Anna	.739	1.000	.172	-1.70	.949
Arif	.063	16.000	.000	1.20	.632
Arthur	.047	39.000	.000	1.90	.316
Astrid	.495	1.500	.084	-1.40	1.265
Avery	.127	6.000	.000	.00	1.054
Ayse	.221	4.118	.001	-.60	1.075
Ben	.048	17.335	.000	1.70	.675
Benjamin	.047	39.000	.000	1.90	.316
Bianca	.982	1.000	.172	-1.90	.316
Bircan	.059	15.461	.000	1.30	.675
Cameron	.074	6.500	.000	.60	1.265
Cengiz	.056	15.377	.000	1.40	.699
Christine	.848	1.964	.041	-1.70	.483
Dace	.100	9.000	.000	.40	.843

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Daniel	.047	39.000	.000	1.90	.316
David	.046	19.000	.000	1.80	.632
Denise	.446	1.769	.055	-1.30	1.252
Deniz	.049	7.686	.000	1.20	1.317
Dennis	.048	17.335	.000	1.70	.675
Dina	.709	1.500	.084	-1.60	.843
Diniz	.174	4.311	.001	-.40	1.174
Dominique	.359	3.354	.004	-1.00	.943
Dylan	.050	9.000	.000	1.30	1.160
Edvin	.048	28.500	.000	1.80	.422
Elena	.982	1.000	.172	-1.90	.316
Elias	.046	11.784	.000	1.60	.966
Elise	.495	1.500	.084	-1.40	1.265
Elvira	.848	1.964	.041	-1.70	.483
Emilia	1.000			-2.00	.000
Emma	.617	1.627	.069	-1.50	.972
Erik	.046			2.00	.000
Erin	.117	2.940	.008	-.30	1.829
Esau	.062	9.000	.000	1.00	1.054
Ethan	.048	17.335	.000	1.70	.675
Eva	.739	1.000	.172	-1.70	.949
Evan	.056	15.377	.000	1.40	.699
Ezra	.174	4.311	.001	-.40	1.174
Fabian	.050	24.222	.000	1.70	.483
Faizah	.247	3.881	.002	-.70	1.059
Felix	.047	39.000	.000	1.90	.316

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Finlay	.056	9.798	.000	1.20	1.033
Finn	.050	24.222	.000	1.70	.483
Fiona	.548	1.246	.122	-1.50	1.269
Florence	.742	1.809	.052	-1.60	.699
Florian	.046	19.000	.000	1.80	.632
Franz	.046			2.00	.000
Freja	.475	3.207	.005	-1.20	.789
Gan	.077	11.225	.000	.80	.789
Gerhard	.046			2.00	.000
Gideon	.050	16.282	.000	1.60	.699
Gil	.093	9.303	.000	.50	.850
Grace	.848	1.964	.041	-1.70	.483
Guilherme	.096	6.273	.000	.30	1.160
Gustav	.046			2.00	.000
Haakon	.050	24.222	.000	1.70	.483
Hamid	.050	24.222	.000	1.70	.483
Hannah	.955	1.500	.084	-1.80	.422
Hoa	.100	9.000	.000	.40	.843
Hugo	.046			2.00	.000
Hung	.086	12.650	.000	.70	.675
Ida	.848	1.964	.041	-1.70	.483
Ingrid	.955	1.500	.084	-1.80	.422
Irfan	.082	6.708	.000	.50	1.179
Isabell	1.000			-2.00	.000
Jade	.306	2.372	.021	-1.00	1.333
Jaime	.085	6.000	.000	.40	1.265

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Jan	.047	39.000	.000	1.90	.316
Jana	1.000			-2.00	.000
Janis	.117	4.146	.001	-.10	1.449
Jemina	.675	2.236	.026	-1.50	.707
Johannes	.046			2.00	.000
Jonas	.048	28.500	.000	1.80	.422
José	.050	16.282	.000	1.60	.699
Josua	.048	28.500	.000	1.80	.422
Julia	1.000			-2.00	.000
Julian	.046	11.784	.000	1.60	.966
Justin	.047	39.000	.000	1.90	.316
Katharina	.982	1.000	.172	-1.90	.316
Keiki	.147	5.014	.000	-.20	1.135
Kim	.178	3.503	.003	-.50	1.354
Klemens	.047	39.000	.000	1.90	.316
Kyle	.054	11.000	.000	1.30	.949
Lana	.982	1.000	.172	-1.90	.316
Lara	1.000			-2.00	.000
Larissa	1.000			-2.00	.000
Lærke	.096	7.856	.000	.40	.966
Latif	.089	8.135	.000	.50	.972
Laura	.955	1.500	.084	-1.80	.422
Lena	.982	1.000	.172	-1.90	.316
Leon	.046	19.000	.000	1.80	.632
Leonie	.982	1.000	.172	-1.90	.316
Leyla	1.000			-2.00	.000

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Liam	.053	8.913	.000	1.20	1.135
Luoanne	.778	2.449	.018	-1.60	.516
Louis	.051	8.232	.000	1.20	1.229
Lovro	.071	14.230	.000	1.00	.667
Lucie	.955	1.500	.084	-1.80	.422
Lucy	.982	1.000	.172	-1.90	.316
Lukas	.046			2.00	.000
Madison	.141	3.073	.007	-.40	1.647
Makani	.136	5.460	.000	-.10	1.101
Mara	.809	1.406	.097	-1.70	.675
Marcel	.050	24.222	.000	1.70	.483
Marek	.050	24.222	.000	1.70	.483
Marie	.955	1.500	.084	-1.80	.422
Marit	.295	2.236	.026	-1.00	1.414
Markus	.046			2.00	.000
Matej	.050	16.282	.000	1.60	.699
Matthias	.046			2.00	.000
Maxime	.156	3.539	.003	-.40	1.430
Maximilian	.046			2.00	.000
Maya	.869	1.000	.172	-1.80	.632
Mercedes	.177	2.806	.010	-.60	1.578
Mia	.982	1.000	.172	-1.90	.316
Michal	.051	11.129	.000	1.40	.966
Mikkel	.046	19.000	.000	1.80	.632
Moritz	.046			2.00	.000
Nadim	.077	7.005	.000	.60	1.174

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Nadine	1.000			-2.00	.000
Natalie	.982	1.000	.172	-1.90	.316
Nathan	.052	22.045	.000	1.60	.516
Nico	.050	16.282	.000	1.60	.699
Nikolaj	.047	39.000	.000	1.90	.316
Nina	1.000			-2.00	.000
Noah	.053	15.652	.000	1.50	.707
Oliver	.047	39.000	.000	1.90	.316
Olivia	.955	1.500	.084	-1.80	.422
Olle	.059	15.461	.000	1.30	.675
Ondrej	.053	15.652	.000	1.50	.707
Osman	.046			2.00	.000
Otto	.047	39.000	.000	1.90	.316
Patrick	.044	12.333	.000	1.70	.949
Pedro	.047	39.000	.000	1.90	.316
Pinar	.117	4.146	.001	-.10	1.449
Pua	.142	6.042	.000	-.10	.994
Pualani	.104	7.667	.000	.30	.949
Rachel	.742	1.809	.052	-1.60	.699
Richard	.046			2.00	.000
Rita	.955	1.500	.084	-1.80	.422
Robin	.060	9.858	.000	1.10	.994
Ronja	.617	1.627	.069	-1.50	.972
Roque	.053	15.652	.000	1.50	.707
Ryan	.048	17.335	.000	1.70	.675
Sabri	.289	4.129	.001	-.80	.919

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Safa	.208	4.881	.000	-.50	.972
Saga	.217	5.582	.000	-.50	.850
Sarah	.982	1.000	.172	-1.90	.316
Sascha	.060	9.858	.000	1.10	.994
Sebastian	.044	12.333	.000	1.70	.949
Seher	.082	6.708	.000	.50	1.179
Selina	.955	1.500	.084	-1.80	.422
Seval	.068	11.619	.000	1.00	.816
Sidney	.197	6.000	.000	-.40	.843
Simon	.047	39.000	.000	1.90	.316
Simone	.464	1.909	.044	-1.30	1.160
Sofia	.982	1.000	.172	-1.90	.316
Sophie	.982	1.000	.172	-1.90	.316
Susanne	.982	1.000	.172	-1.90	.316
Sven	.046			2.00	.000
Tahir	.059	15.461	.000	1.30	.675
Tanja	.739	1.000	.172	-1.70	.949
Tarek	.048	13.500	.000	1.60	.843
Thao	.062	11.196	.000	1.10	.876
Tim	.048	28.500	.000	1.80	.422
Tobias	.047	39.000	.000	1.90	.316
Tom	.047	39.000	.000	1.90	.316
Tung	.077	11.225	.000	.80	.789
Tuva	.127	6.000	.000	.00	1.054
Vega	.096	6.273	.000	.30	1.160
Vera	.982	1.000	.172	-1.90	.316

(Table B-3.2 continues)

(Table B-3.2 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Vilde	.133	6.708	.000	.00	.943
Wedat	.077	11.225	.000	.80	.789
William	.044	12.333	.000	1.70	.949
Yael	.087	9.750	.000	.60	.843
Yanis	.055	7.115	.000	1.00	1.333
Yasar	.065	13.286	.000	1.10	.738
Zada	.328	3.973	.002	-.90	.876

Table B-3.3

Achievement levels and t-values for $x_p \equiv \mu_0 = +2.00$ (perfect male mean).

	A(x,y)	t	Sig. (1-tailed)	M	SD
Aaron	.982	-1.000	.172	1.90	.316
Abigail	.050	-24.222	.000	-1.70	.483
Adam	.982	-1.000	.172	1.90	.316
Alani	.079	-6.228	.000	-.50	1.269
Alexander	.955	-1.500	.084	1.80	.422
Alexandra	.046	-10.247	.000	-1.50	1.080
Alina	.046	-19.000	.000	-1.80	.632
Amelie	.047	-39.000	.000	-1.90	.316
Andreas	.679	-1.309	.111	1.60	.966
Anna	.044	-12.333	.000	-1.70	.949
Arif	.498	-4.000	.002	1.20	.632
Arthur	.982	-1.000	.172	1.90	.316
Astrid	.045	-8.500	.000	-1.40	1.265
Avery	.127	-6.000	.000	.00	1.054
Ayse	.080	-7.649	.000	-.60	1.075
Ben	.809	-1.406	.097	1.70	.675
Benjamin	.982	-1.000	.172	1.90	.316
Bianca	.047	-39.000	.000	-1.90	.316
Bircan	.551	-3.280	.005	1.30	.675
Cameron	.204	-3.500	.003	.60	1.265
Cengiz	.610	-2.714	.012	1.40	.699
Christine	.050	-24.222	.000	-1.70	.483
Dace	.197	-6.000	.000	.40	.843
Daniel	.982	-1.000	.172	1.90	.316
David	.869	-1.000	.172	1.80	.632

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Denise	.048	-8.337	.000	-1.30	1.252
Deniz	.387	-1.922	.043	1.20	1.317
Dennis	.809	-1.406	.097	1.70	.675
Dina	.048	-13.500	.000	-1.60	.843
Diniz	.089	-6.466	.000	-.40	1.174
Dominique	.065	-10.062	.000	-1.00	.943
Dylan	.464	-1.909	.044	1.30	1.160
Edvin	.955	-1.500	.084	1.80	.422
Elena	.047	-39.000	.000	-1.90	.316
Elias	.679	-1.309	.111	1.60	.966
Elise	.045	-8.500	.000	-1.40	1.265
Elvira	.050	-24.222	.000	-1.70	.483
Emilia	.046			-2.00	.000
Emma	.048	-11.389	.000	-1.50	.972
Erik	1.000			2.00	.000
Erin	.071	-3.977	.002	-.30	1.829
Esau	.344	-3.000	.007	1.00	1.054
Ethan	.809	-1.406	.097	1.70	.675
Eva	.044	-12.333	.000	-1.70	.949
Evan	.610	-2.714	.012	1.40	.699
Ezra	.089	-6.466	.000	-.40	1.174
Fabian	.848	-1.964	.041	1.70	.483
Faizah	.075	-8.060	.000	-.70	1.059
Felix	.982	-1.000	.172	1.90	.316
Finlay	.436	-2.449	.018	1.20	1.033
Finn	.848	-1.964	.041	1.70	.483

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Fiona	.043	-8.720	.000	-1.50	1.269
Florence	.050	-16.282	.000	-1.60	.699
Florian	.869	-1.000	.172	1.80	.632
Franz	1.000			2.00	.000
Freja	.061	-12.829	.000	-1.20	.789
Gan	.303	-4.811	.000	.80	.789
Gerhard	1.000			2.00	.000
Gideon	.742	-1.809	.052	1.60	.699
Gil	.217	-5.582	.000	.50	.850
Grace	.050	-24.222	.000	-1.70	.483
Guilherme	.159	-4.636	.001	.30	1.160
Gustav	1.000			2.00	.000
Haakon	.848	-1.964	.041	1.70	.483
Hamid	.848	-1.964	.041	1.70	.483
Hannah	.048	-28.500	.000	-1.80	.422
Hoa	.197	-6.000	.000	.40	.843
Hugo	1.000			2.00	.000
Hung	.281	-6.091	.000	.70	.675
Ida	.050	-24.222	.000	-1.70	.483
Ingrid	.048	-28.500	.000	-1.80	.422
Irfan	.191	-4.025	.001	.50	1.179
Isabell	.046			-2.00	.000
Jade	.055	-7.115	.000	-1.00	1.333
Jaime	.167	-4.000	.002	.40	1.265
Jan	.982	-1.000	.172	1.90	.316
Jana	.046			-2.00	.000

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Janis	.099	-4.583	.001	-.10	1.449
Jemina	.053	-15.652	.000	-1.50	.707
Johannes	1.000			2.00	.000
Jonas	.955	-1.500	.084	1.80	.422
José	.742	-1.809	.052	1.60	.699
Josua	.955	-1.500	.084	1.80	.422
Julia	.046			-2.00	.000
Julian	.679	-1.309	.111	1.60	.966
Justin	.982	-1.000	.172	1.90	.316
Katharina	.047	-39.000	.000	-1.90	.316
Keiki	.105	-6.128	.000	-.20	1.135
Kim	.076	-5.839	.000	-.50	1.354
Klemens	.982	-1.000	.172	1.90	.316
Kyle	.503	-2.333	.022	1.30	.949
Lana	.047	-39.000	.000	-1.90	.316
Lara	.046			-2.00	.000
Larissa	.046			-2.00	.000
Lærke	.189	-5.237	.000	.40	.966
Latif	.208	-4.881	.000	.50	.972
Laura	.048	-28.500	.000	-1.80	.422
Lena	.047	-39.000	.000	-1.90	.316
Leon	.869	-1.000	.172	1.80	.632
Leonie	.047	-39.000	.000	-1.90	.316
Leyla	.046			-2.00	.000
Liam	.418	-2.228	.026	1.20	1.135
Luoanne	.052	-22.045	.000	-1.60	.516

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Louis	.402	-2.058	.035	1.20	1.229
Lovro	.392	-4.743	.001	1.00	.667
Lucie	.048	-28.500	.000	-1.80	.422
Lucy	.047	-39.000	.000	-1.90	.316
Lukas	1.000			2.00	.000
Madison	.072	-4.609	.001	-.40	1.647
Makani	.115	-6.034	.000	-.10	1.101
Mara	.048	-17.335	.000	-1.70	.675
Marcel	.848	-1.964	.041	1.70	.483
Marek	.848	-1.964	.041	1.70	.483
Marie	.048	-28.500	.000	-1.80	.422
Marit	.053	-6.708	.000	-1.00	1.414
Markus	1.000			2.00	.000
Matej	.742	-1.809	.052	1.60	.699
Matthias	1.000			2.00	.000
Maxime	.079	-5.308	.000	-.40	1.430
Maximilian	1.000			2.00	.000
Maya	.046	-19.000	.000	-1.80	.632
Mercedes	.064	-5.212	.000	-.60	1.578
Mia	.047	-39.000	.000	-1.90	.316
Michal	.558	-1.964	.041	1.40	.966
Mikkel	.869	-1.000	.172	1.80	.632
Moritz	1.000			2.00	.000
Nadim	.212	-3.772	.002	.60	1.174
Nadine	.046			-2.00	.000
Natalie	.047	-39.000	.000	-1.90	.316

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Nathan	.778	-2.449	.018	1.60	.516
Nico	.742	-1.809	.052	1.60	.699
Nikolaj	.982	-1.000	.172	1.90	.316
Nina	.046			-2.00	.000
Noah	.675	-2.236	.026	1.50	.707
Oliver	.982	-1.000	.172	1.90	.316
Olivia	.048	-28.500	.000	-1.80	.422
Olle	.551	-3.280	.005	1.30	.675
Ondrej	.675	-2.236	.026	1.50	.707
Osman	1.000			2.00	.000
Otto	.982	-1.000	.172	1.90	.316
Patrick	.739	-1.000	.172	1.70	.949
Pedro	.982	-1.000	.172	1.90	.316
Pinar	.099	-4.583	.001	-.10	1.449
Pua	.120	-6.678	.000	-.10	.994
Pualani	.173	-5.667	.000	.30	.949
Rachel	.050	-16.282	.000	-1.60	.699
Richard	1.000			2.00	.000
Rita	.048	-28.500	.000	-1.80	.422
Robin	.394	-2.862	.009	1.10	.994
Ronja	.048	-11.389	.000	-1.50	.972
Roque	.675	-2.236	.026	1.50	.707
Ryan	.809	-1.406	.097	1.70	.675
Sabri	.074	-9.635	.000	-.80	.919
Safa	.089	-8.135	.000	-.50	.972
Saga	.093	-9.303	.000	-.50	.850

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
Sarah	.047	-39.000	.000	-1.90	.316
Sascha	.394	-2.862	.009	1.10	.994
Sebastian	.739	-1.000	.172	1.70	.949
Seher	.191	-4.025	.001	.50	1.179
Selina	.048	-28.500	.000	-1.80	.422
Seval	.375	-3.873	.002	1.00	.816
Sidney	.100	-9.000	.000	-.40	.843
Simon	.982	-1.000	.172	1.90	.316
Simone	.050	-9.000	.000	-1.30	1.160
Sofia	.047	-39.000	.000	-1.90	.316
Sophie	.047	-39.000	.000	-1.90	.316
Susanne	.047	-39.000	.000	-1.90	.316
Sven	1.000			2.00	.000
Tahir	.551	-3.280	.005	1.30	.675
Tanja	.044	-12.333	.000	-1.70	.949
Tarek	.709	-1.500	.084	1.60	.843
Thao	.412	-3.250	.005	1.10	.876
Tim	.955	-1.500	.168	1.80	.422
Tobias	.982	-1.000	.172	1.90	.316
Tom	.982	-1.000	.172	1.90	.316
Tung	.303	-4.811	.000	.80	.789
Tuva	.127	-6.000	.000	.00	1.054
Vega	.159	-4.636	.001	.30	1.160
Vera	.047	-39.000	.000	-1.90	.316
Vilde	.133	-6.708	.000	.00	.943
Wedat	.303	-4.811	.000	.80	.789

(Table B-3.3 continues)

(Table B-3.3 continued)

	A(x,y)	t	Sig. (1-tailed)	M	SD
William	.739	-1.000	.172	1.70	.949
Yael	.240	-5.250	.000	.60	.843
Yanis	.306	-2.372	.021	1.00	1.333
Yasar	.431	-3.857	.002	1.10	.738
Zada	.071	-10.474	.000	-.90	.876

Table B-3.4

Achievement levels and t-values for $x_p \equiv \mu_0 = \pm.00$ (perfect neutral mean).

	A(x,y)	t	Sig. (2-tailed)	M	SD
Aaron	.171	19.000	.000	1.90	.316
Abigail	.198	-11.129	.000	-1.70	.483
Adam	.171	19.000	.000	1.90	.316
Alani	.548	-1.246	.244	-.50	1.269
Alexander	.183	13.500	.000	1.80	.422
Alexandra	.199	-4.392	.002	-1.50	1.080
Alina	.174	-9.000	.000	-1.80	.632
Amelie	.171	-19.000	.000	-1.90	.316
Andreas	.189	5.237	.001	1.60	.966
Anna	.173	-5.667	.000	-1.70	.949
Arif	.317	6.000	.000	1.20	.632
Arthur	.171	19.000	.000	1.90	.316
Astrid	.204	-3.500	.007	-1.40	1.265
Avery	.794	.000	1.000	.00	1.054
Ayse	.535	-1.765	.111	-.60	1.075
Ben	.189	7.965	.000	1.70	.675
Benjamin	.171	19.000	.000	1.90	.316
Bianca	.171	-19.000	.000	-1.90	.316
Bircan	.281	6.091	.000	1.30	.675
Cameron	.495	1.500	.168	.60	1.265
Cengiz	.252	6.332	.000	1.40	.699
Christine	.198	-11.129	.000	-1.70	.483
Dace	.709	1.500	.168	.40	.843
Daniel	.171	19.000	.000	1.90	.316
David	.174	9.000	.000	1.80	.632

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
Denise	.228	-3.284	.009	-1.30	1.252
Deniz	.247	2.882	.018	1.20	1.317
Dennis	.189	7.965	.000	1.70	.675
Dina	.197	-6.000	.000	-1.60	.843
Diniz	.626	-1.078	.309	-.40	1.174
Dominique	.359	-3.354	.008	-1.00	.943
Dylan	.237	3.545	.006	1.30	1.160
Edvin	.183	13.500	.000	1.80	.422
Elena	.171	-19.000	.000	-1.90	.316
Elias	.189	5.237	.001	1.60	.966
Elise	.204	-3.500	.007	-1.40	1.265
Elvira	.198	-11.129	.000	-1.70	.483
Emilia	.160			-2.00	.000
Emma	.208	-4.881	.001	-1.50	.972
Erik	.160			2.00	.000
Erin	.503	-.519	.616	-.30	1.829
Esau	.344	3.000	.015	1.00	1.054
Ethan	.189	7.965	.000	1.70	.675
Eva	.173	-5.667	.000	-1.70	.949
Evan	.252	6.332	.000	1.40	.699
Ezra	.626	-1.078	.309	-.40	1.174
Fabian	.198	11.129	.000	1.70	.483
Faizah	.483	-2.090	.066	-.70	1.059
Felix	.171	19.000	.000	1.90	.316
Finlay	.277	3.674	.005	1.20	1.033
Finn	.198	11.129	.000	1.70	.483

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
Fiona	.184	-3.737	.005	-1.50	1.269
Florence	.206	-7.236	.000	-1.60	.699
Florian	.174	9.000	.000	1.80	.632
Franz	.160			2.00	.000
Freja	.303	-4.811	.001	-1.20	.789
Gan	.475	3.207	.011	.80	.789
Gerhard	.160			2.00	.000
Gideon	.206	7.236	.000	1.60	.699
Gil	.645	1.861	.096	.50	.850
Grace	.198	-11.129	.000	-1.70	.483
Guilherme	.681	.818	.434	.30	1.160
Gustav	.160			2.00	.000
Haakon	.198	11.129	.000	1.70	.483
Hamid	.198	11.129	.000	1.70	.483
Hannah	.183	-13.500	.000	-1.80	.422
Hoa	.709	1.500	.168	.40	.843
Hugo	.160			2.00	.000
Hung	.551	3.280	.010	.70	.675
Ida	.198	-11.129	.000	-1.70	.483
Ingrid	.183	-13.500	.000	-1.80	.422
Irfan	.569	1.342	.213	.50	1.179
Isabell	.160			-2.00	.000
Jade	.306	-2.372	.042	-1.00	1.333
Jaime	.602	1.000	.343	.40	1.265
Jan	.171	19.000	.000	1.90	.316
Jana	.160			-2.00	.000

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
Janis	.663	-.218	.832	-.10	1.449
Jemina	.227	-6.708	.000	-1.50	.707
Johannes	.160			2.00	.000
Jonas	.183	13.500	.000	1.80	.422
José	.206	7.236	.000	1.60	.699
Josua	.183	13.500	.000	1.80	.422
Julia	.160			-2.00	.000
Julian	.189	5.237	.001	1.60	.966
Justin	.171	19.000	.000	1.90	.316
Katharina	.171	-19.000	.000	-1.90	.316
Keiki	.731	-.557	.591	-.20	1.135
Kim	.528	-1.168	.273	-.50	1.354
Klemens	.171	19.000	.000	1.90	.316
Kyle	.257	4.333	.002	1.30	.949
Lana	.171	-19.000	.000	-1.90	.316
Lara	.160			-2.00	.000
Larissa	.160			-2.00	.000
Lærke	.679	1.309	.223	.40	.966
Latif	.617	1.627	.138	.50	.972
Laura	.183	-13.500	.000	-1.80	.422
Lena	.171	-19.000	.000	-1.90	.316
Leon	.174	9.000	.000	1.80	.632
Leonie	.171	-19.000	.000	-1.90	.316
Leyla	.160			-2.00	.000
Liam	.266	3.343	.009	1.20	1.135
Luoanne	.216	-9.798	.000	-1.60	.516

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
Louis	.256	3.087	.013	1.20	1.229
Lovro	.392	4.743	.001	1.00	.667
Lucie	.183	-13.500	.000	-1.80	.422
Lucy	.171	-19.000	.000	-1.90	.316
Lukas	.160			2.00	.000
Madison	.507	-.768	.462	-.40	1.647
Makani	.770	-.287	.780	-.10	1.101
Mara	.189	-7.965	.000	-1.70	.675
Marcel	.198	11.129	.000	1.70	.483
Marek	.198	11.129	.000	1.70	.483
Marie	.183	-13.500	.000	-1.80	.422
Marit	.295	-2.236	.052	-1.00	1.414
Markus	.160			2.00	.000
Matej	.206	7.236	.000	1.60	.699
Matthias	.160			2.00	.000
Maxime	.560	-.885	.399	-.40	1.430
Maximilian	.160			2.00	.000
Maya	.174	-9.000	.000	-1.80	.632
Mercedes	.430	-1.203	.260	-.60	1.578
Mia	.171	-19.000	.000	-1.90	.316
Michal	.230	4.583	.001	1.40	.966
Mikkel	.174	9.000	.000	1.80	.632
Moritz	.160			2.00	.000
Nadim	.514	1.616	.140	.60	1.174
Nadine	.160			-2.00	.000
Natalie	.171	-19.000	.000	-1.90	.316

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
Nathan	.216	9.798	.000	1.60	.516
Nico	.206	7.236	.000	1.60	.699
Nikolaj	.171	19.000	.000	1.90	.316
Nina	.160			-2.00	.000
Noah	.227	6.708	.000	1.50	.707
Oliver	.171	19.000	.000	1.90	.316
Olivia	.183	-13.500	.000	-1.80	.422
Olle	.281	6.091	.000	1.30	.675
Ondrej	.227	6.708	.000	1.50	.707
Osman	.160			2.00	.000
Otto	.171	19.000	.000	1.90	.316
Patrick	.173	5.667	.000	1.70	.949
Pedro	.171	19.000	.000	1.90	.316
Pinar	.663	-.218	.832	-.10	1.449
Pua	.802	-.318	.758	-.10	.994
Pualani	.739	1.000	.343	.30	.949
Rachel	.206	-7.236	.000	-1.60	.699
Richard	.160			2.00	.000
Rita	.183	-13.500	.000	-1.80	.422
Robin	.314	3.498	.007	1.10	.994
Ronja	.208	-4.881	.001	-1.50	.972
Roque	.227	6.708	.000	1.50	.707
Ryan	.189	7.965	.000	1.70	.675
Sabri	.455	-2.753	.022	-.80	.919
Safa	.617	-1.627	.138	-.50	.972
Saga	.645	-1.861	.096	-.50	.850

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
Sarah	.171	-19.000	.000	-1.90	.316
Sascha	.314	3.498	.007	1.10	.994
Sebastian	.173	5.667	.000	1.70	.949
Seher	.569	1.342	.213	.50	1.179
Selina	.183	-13.500	.000	-1.80	.422
Seval	.375	3.873	.004	1.00	.816
Sidney	.709	-1.500	.168	-.40	.843
Simon	.171	19.000	.000	1.90	.316
Simone	.237	-3.545	.006	-1.30	1.160
Sofia	.171	-19.000	.000	-1.90	.316
Sophie	.171	-19.000	.000	-1.90	.316
Susanne	.171	-19.000	.000	-1.90	.316
Sven	.160			2.00	.000
Tahir	.281	6.091	.000	1.30	.675
Tanja	.173	-5.667	.000	-1.70	.949
Tarek	.197	6.000	.000	1.60	.843
Thao	.328	3.973	.003	1.10	.876
Tim	.183	13.500	.000	1.80	.422
Tobias	.171	19.000	.000	1.90	.316
Tom	.171	19.000	.000	1.90	.316
Tung	.475	3.207	.011	.80	.789
Tuva	.794	.000	1.000	.00	1.054
Vega	.681	.818	.434	.30	1.160
Vera	.171	-19.000	.000	-1.90	.316
Vilde	.828	.000	1.000	.00	.943
Wedat	.475	3.207	.011	.80	.789

(Table B-3.4 continues)

(Table B-3.4 continued)

	A(x,y)	t	Sig. (2-tailed)	M	SD
William	.173	5.667	.000	1.70	.949
Yael	.583	2.250	.051	.60	.843
Yanis	.306	2.372	.042	1.00	1.333
Yasar	.343	4.714	.001	1.10	.738
Zada	.412	-3.250	.010	-.90	.876

Appendix B4 [Pre-Experiment II – Figures]

The figures in appendix B contain all figures used in the paper as well as additional figures referred to in the paper in the pre-experiment II section. All figures are digitally available in the Portable Network Graphic format [*.png] on the appendix CD [/content/files/appendix/02_pre-experiment2/figures/..].

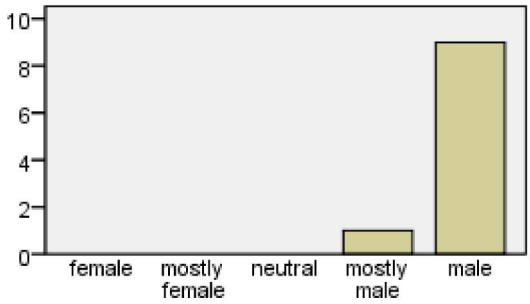


Figure B-4.1. Absolute number of ratings in each category for 'Aaron'

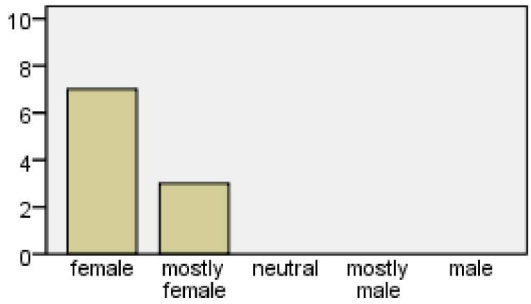


Figure B-4.2. Absolute number of ratings in each category for 'Abigail'

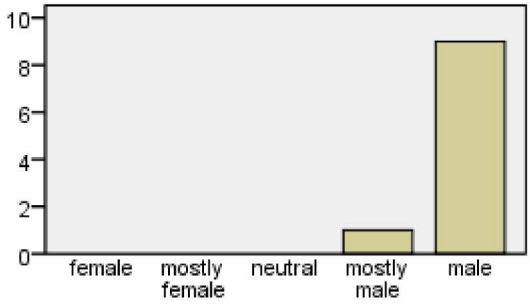


Figure B-4.3. Absolute number of ratings in each category for 'Adam'

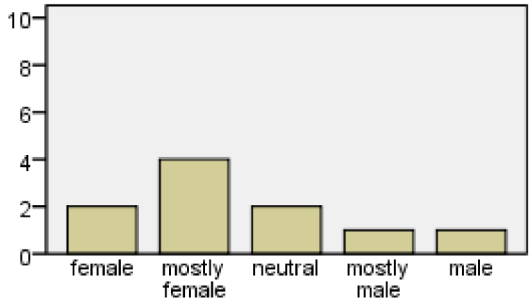


Figure B-4.4. Absolute number of ratings in each category for 'Alani'

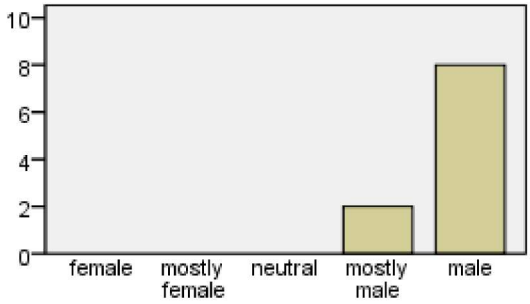


Figure B-4.5. Absolute number of ratings in each category for 'Alexander'

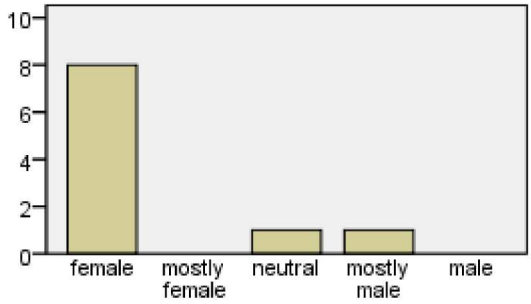


Figure B-4.6. Absolute number of ratings in each category for 'Alexandra'

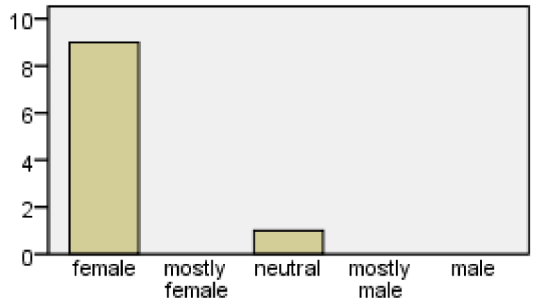


Figure B-4.7. Absolute number of ratings in each category for 'Alina'

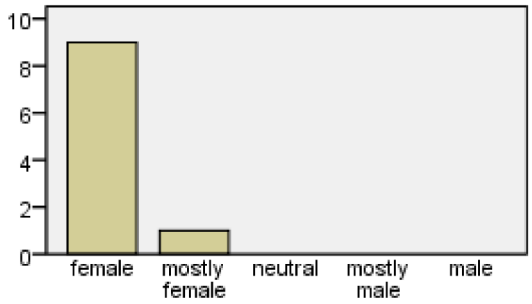


Figure B-4.8. Absolute number of ratings in each category for 'Amelie'

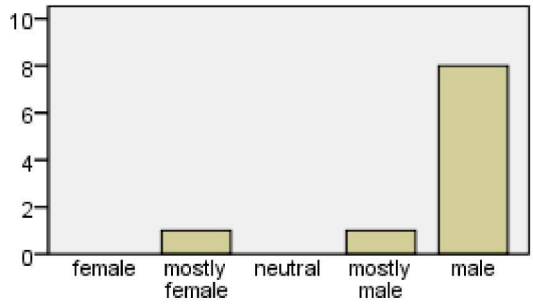


Figure B-4.9. Absolute number of ratings in each category for 'Andreas'

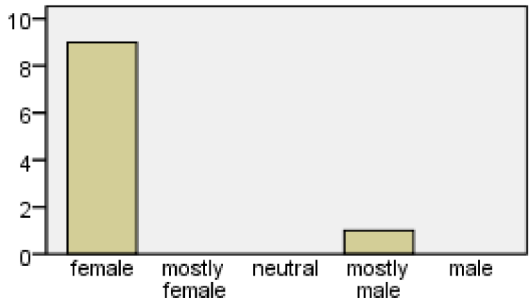


Figure B-4.10. Absolute number of ratings in each category for 'Anna'

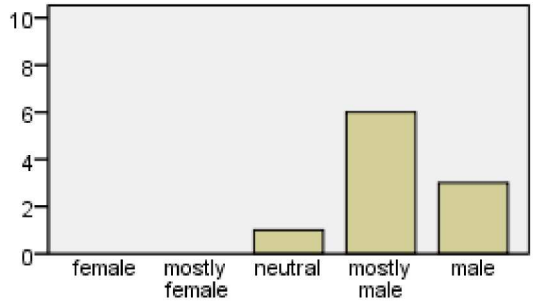


Figure B-4.11. Absolute number of ratings in each category for 'Arif'

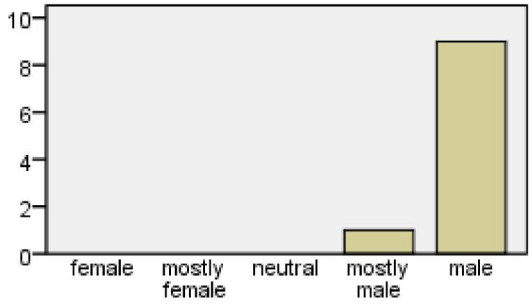


Figure B-4.12. Absolute number of ratings in each category for 'Arthur'

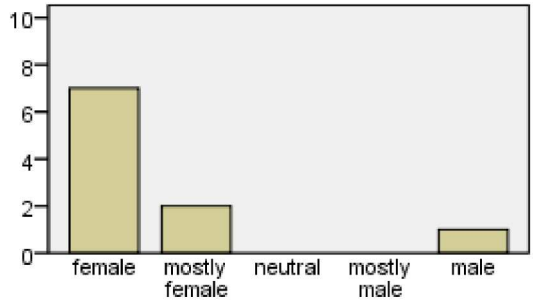


Figure B-4.13. Absolute number of ratings in each category for 'Astrid'

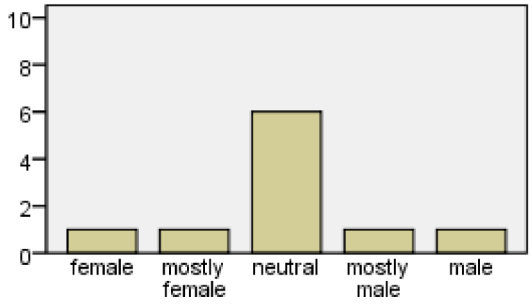


Figure B-4.14. Absolute number of ratings in each category for 'Avery'

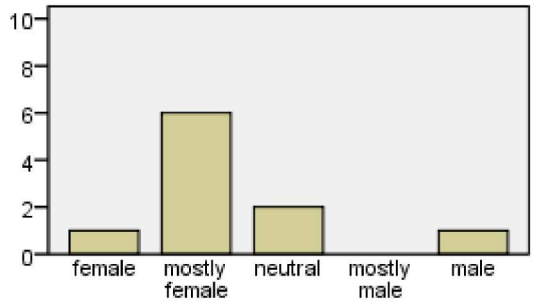


Figure B-4.15. Absolute number of ratings in each category for 'Ayse'

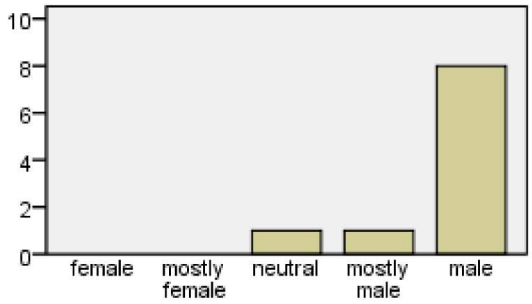


Figure B-4.16. Absolute number of ratings in each category for 'Ben'

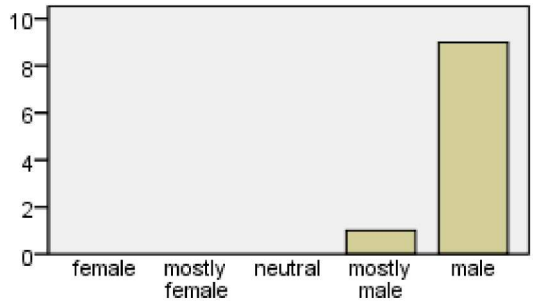


Figure B-4.17. Absolute number of ratings in each category for 'Benjamin'

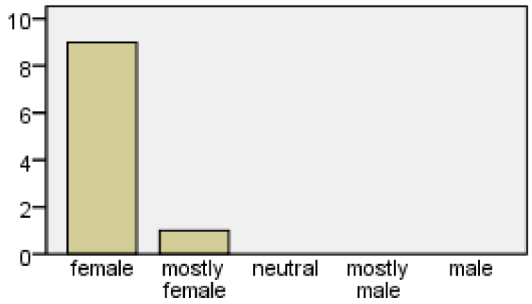


Figure B-4.18. Absolute number of ratings in each category for 'Bianca'

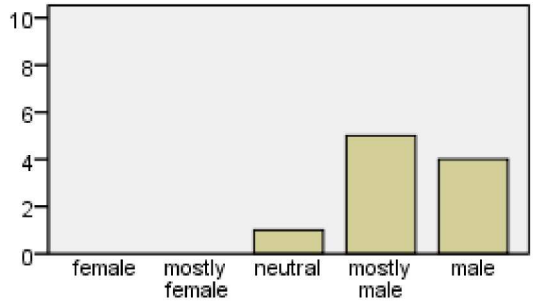


Figure B-4.19. Absolute number of ratings in each category for 'Bircan'

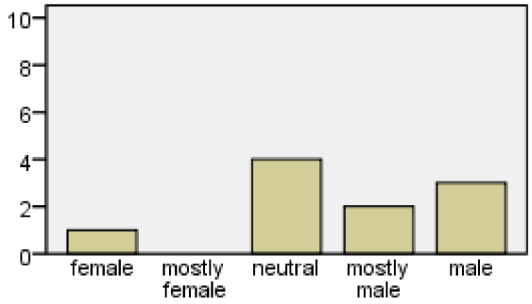


Figure B-4.20. Absolute number of ratings in each category for 'Cameron'

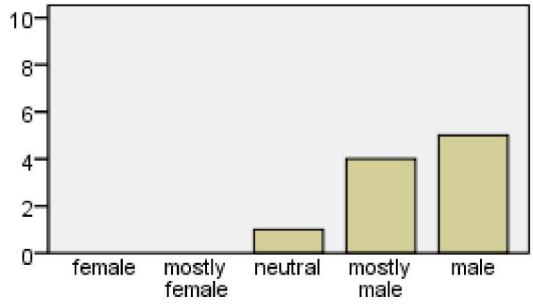


Figure B-4.21. Absolute number of ratings in each category for 'Cengiz'

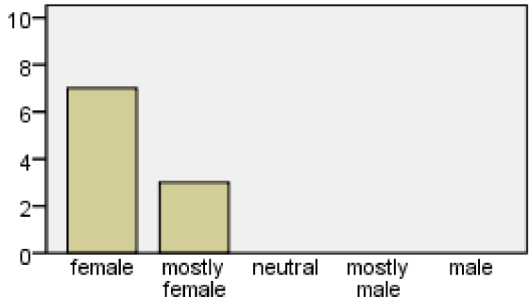


Figure B-4.22. Absolute number of ratings in each category for 'Christine'

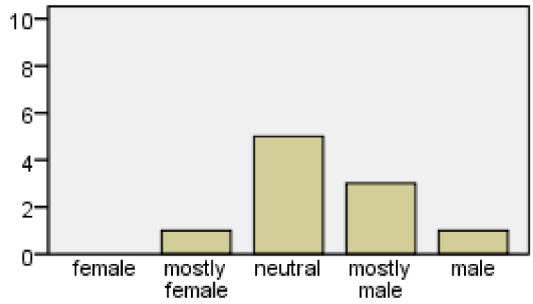


Figure B-4.23. Absolute number of ratings in each category for 'Dace'

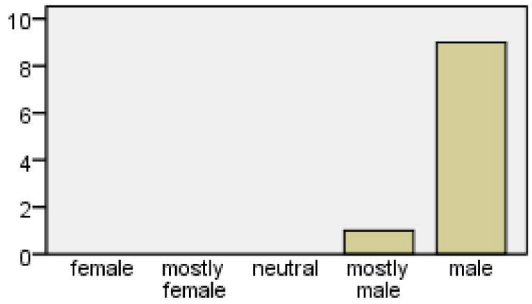


Figure B-4.24. Absolute number of ratings in each category for 'Daniel'

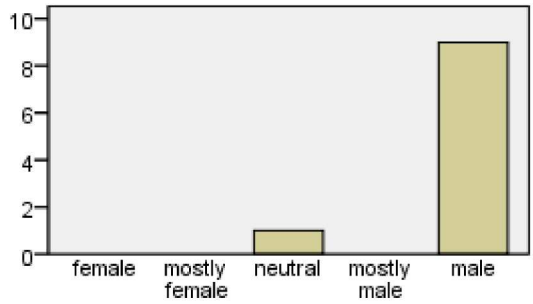


Figure B-4.25. Absolute number of ratings in each category for 'David'

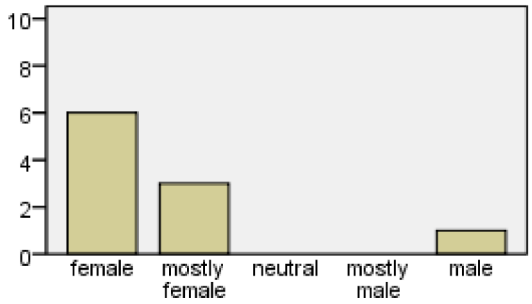


Figure B-4.26. Absolute number of ratings in each category for 'Denise'

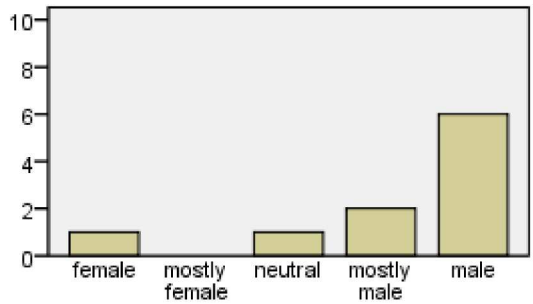


Figure B-4.27. Absolute number of ratings in each category for 'Deniz'

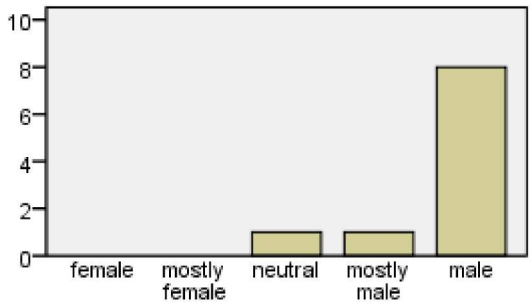


Figure B-4.28. Absolute number of ratings in each category for 'Dennis'

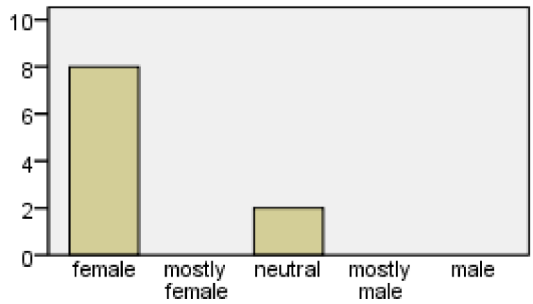


Figure B-4.29. Absolute number of ratings in each category for 'Dina'

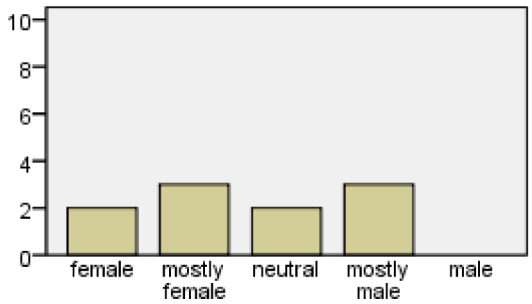


Figure B-4.30. Absolute number of ratings in each category for 'Diniz'

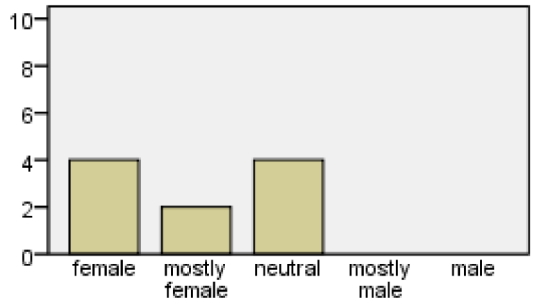


Figure B-4.31. Absolute number of ratings in each category for 'Dominique'

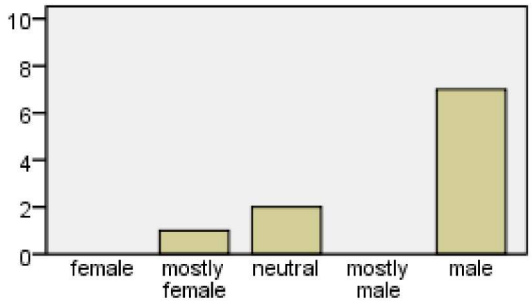


Figure B-4.32. Absolute number of ratings in each category for 'Dylan'

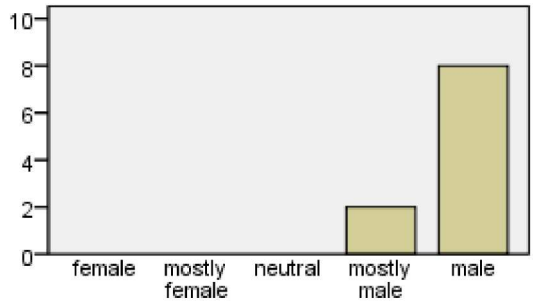


Figure B-4.33. Absolute number of ratings in each category for 'Edvin'

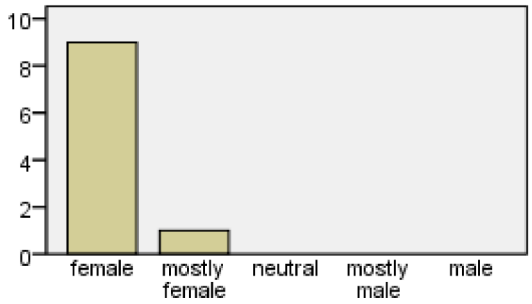


Figure B-4.34. Absolute number of ratings in each category for 'Elena'

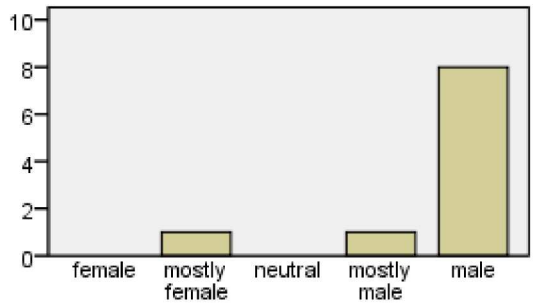


Figure B-4.35. Absolute number of ratings in each category for 'Elias'

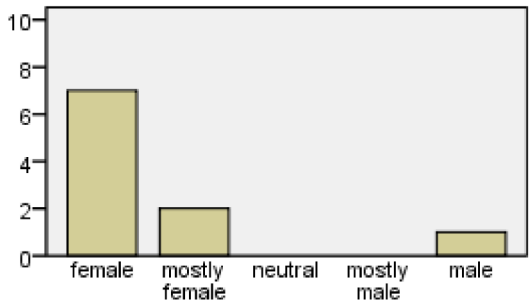


Figure B-4.36. Absolute number of ratings in each category for 'Elise'

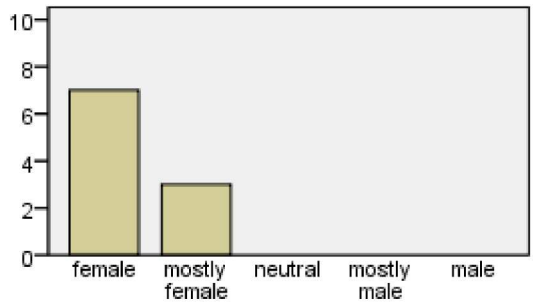


Figure B-4.37. Absolute number of ratings in each category for 'Elvira'

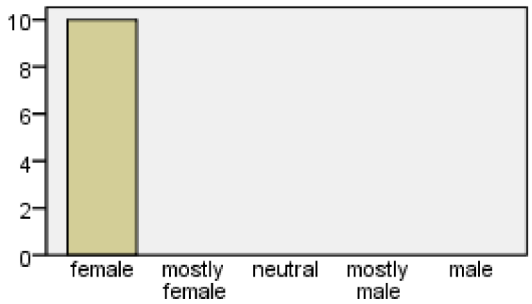


Figure B-4.38. Absolute number of ratings in each category for 'Emilia'

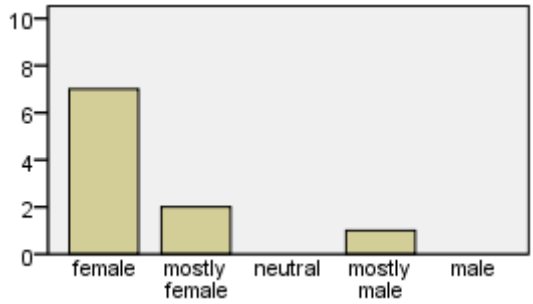


Figure B-4.39. Absolute number of ratings in each category for 'Emma'

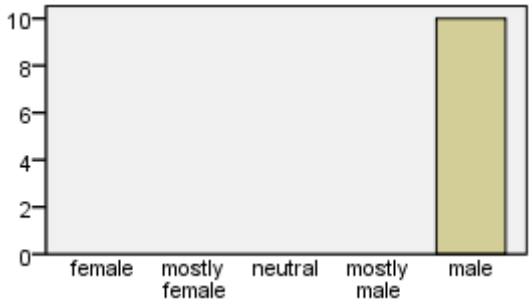


Figure B-4.40. Absolute number of ratings in each category for 'Erik'

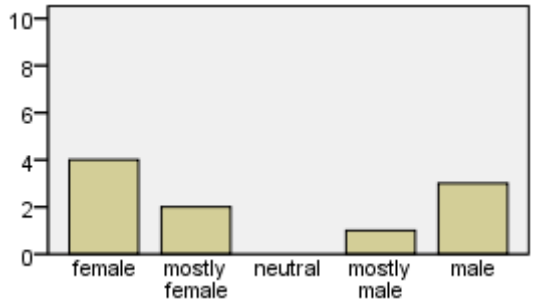


Figure B-4.41. Absolute number of ratings in each category for 'Erin'

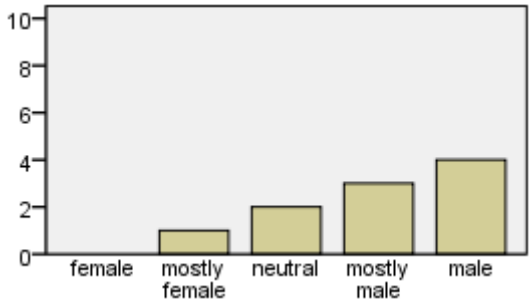


Figure B-4.42. Absolute number of ratings in each category for 'Esau'

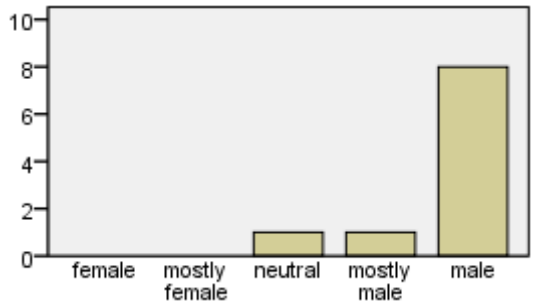


Figure B-4.43. Absolute number of ratings in each category for 'Ethan'

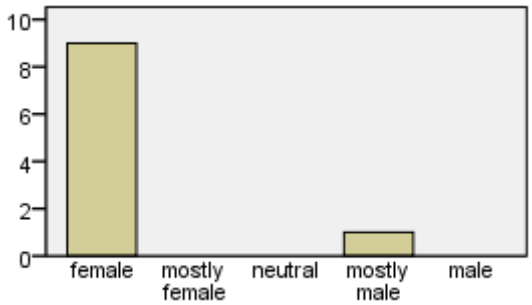


Figure B-4.44. Absolute number of ratings in each category for 'Eva'

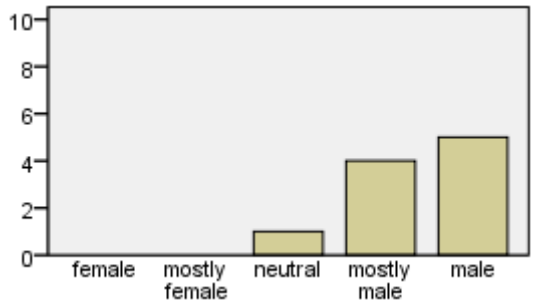


Figure B-4.45. Absolute number of ratings in each category for 'Evan'

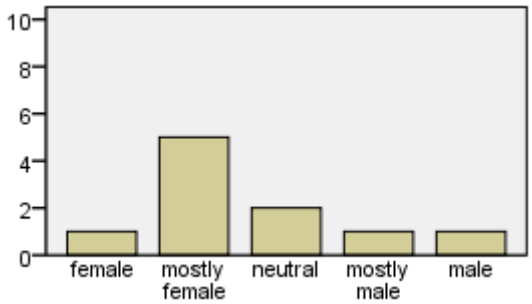


Figure B-4.46. Absolute number of ratings in each category for 'Ezra'

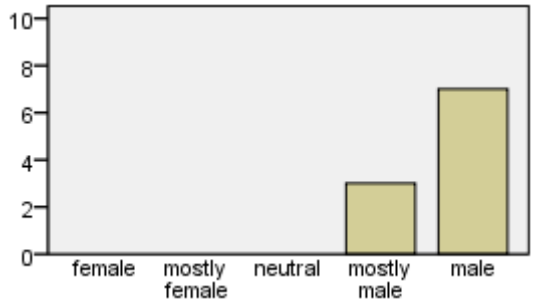


Figure B-4.47. Absolute number of ratings in each category for 'Fabian'

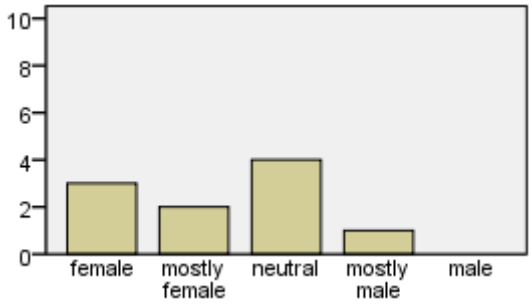


Figure B-4.48. Absolute number of ratings in each category for 'Faizah'

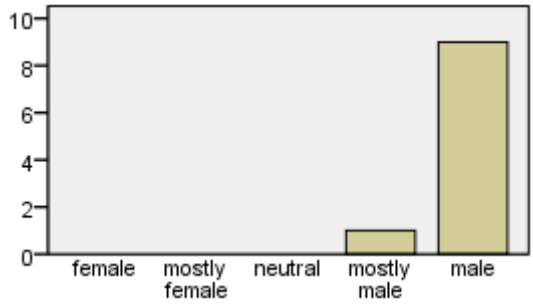


Figure B-4.49. Absolute number of ratings in each category for 'Felix'

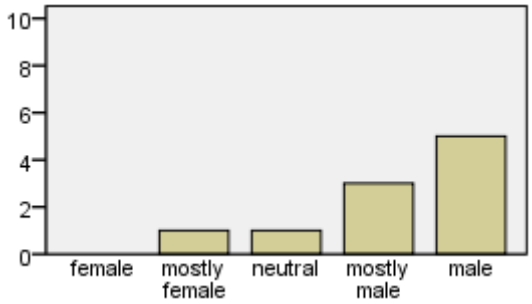


Figure B-4.50. Absolute number of ratings in each category for 'Finlay'

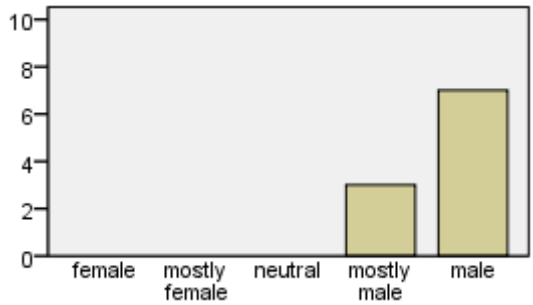


Figure B-4.51. Absolute number of ratings in each category for 'Finn'

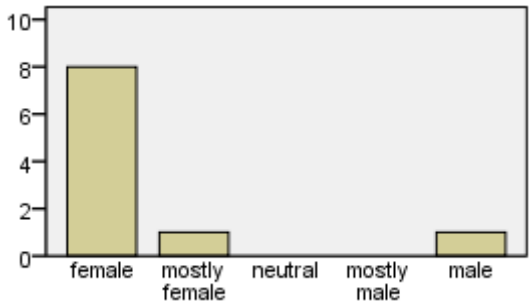


Figure B-4.52. Absolute number of ratings in each category for 'Fiona'

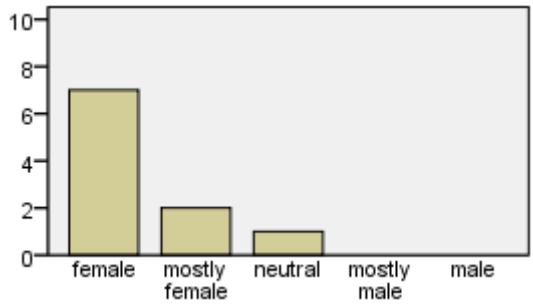


Figure B-4.53. Absolute number of ratings in each category for 'Florence'

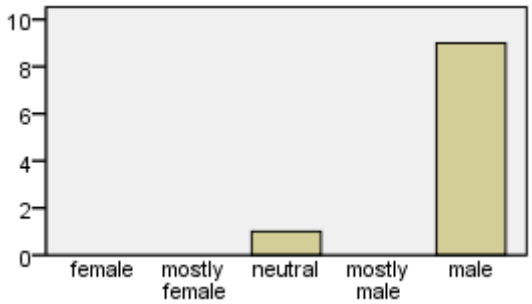


Figure B-4.54. Absolute number of ratings in each category for 'Florian'

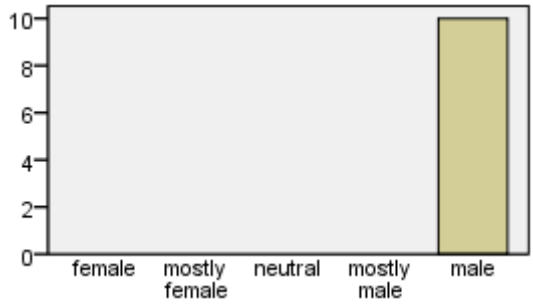


Figure B-4.55. Absolute number of ratings in each category for 'Franz'

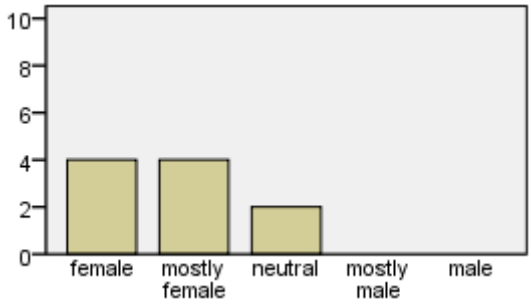


Figure B-4.56. Absolute number of ratings in each category for 'Freja'

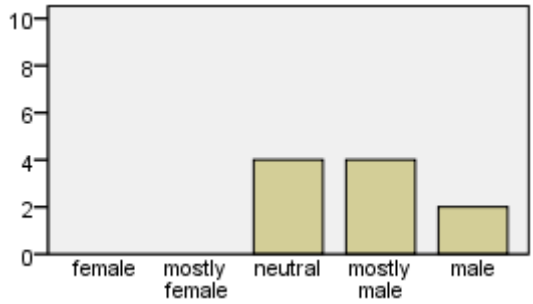


Figure B-4.57. Absolute number of ratings in each category for 'Gan'

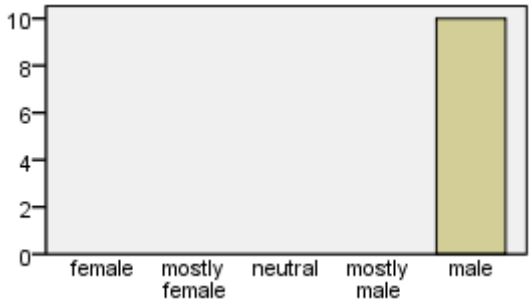


Figure B-4.58. Absolute number of ratings in each category for 'Gerhard'

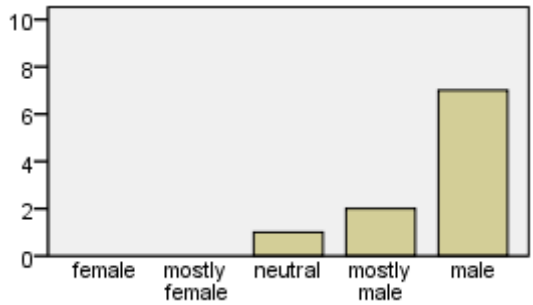


Figure B-4.59. Absolute number of ratings in each category for 'Gideon'

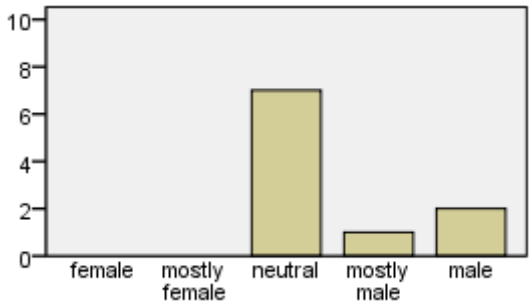


Figure B-4.60. Absolute number of ratings in each category for 'Gil'

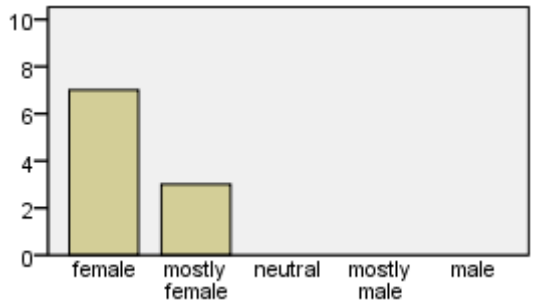


Figure B-4.61. Absolute number of ratings in each category for 'Grace'

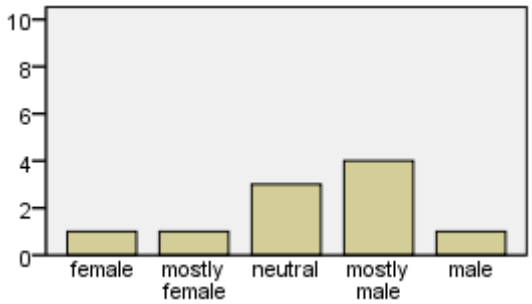


Figure B-4.62. Absolute number of ratings in each category for 'Guilherme'

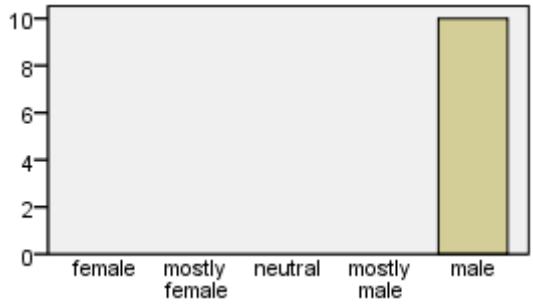


Figure B-4.63. Absolute number of ratings in each category for 'Gustav'

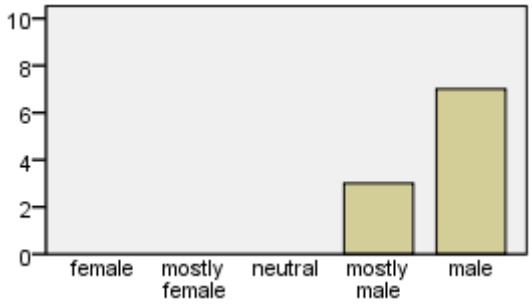


Figure B-4.64. Absolute number of ratings in each category for 'Haakon'

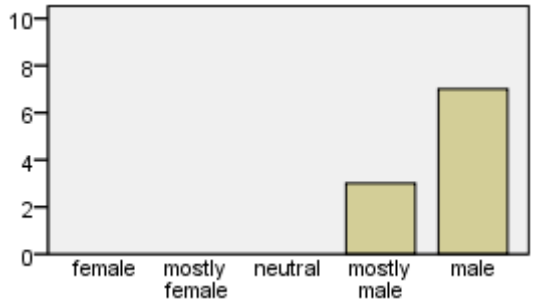


Figure B-4.65. Absolute number of ratings in each category for 'Hamid'

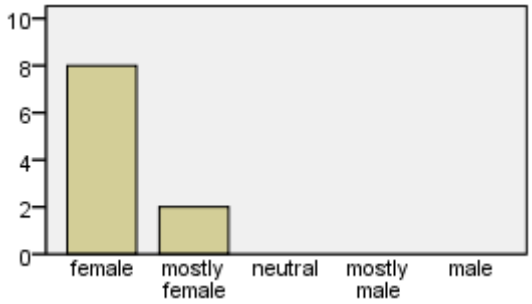


Figure B-4.66. Absolute number of ratings in each category for 'Hannah'

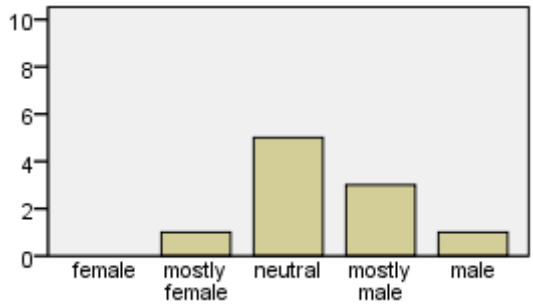


Figure B-4.67. Absolute number of ratings in each category for 'Hoa'

Figure B-4.68. Absolute number of ratings in each category for 'Hugo'

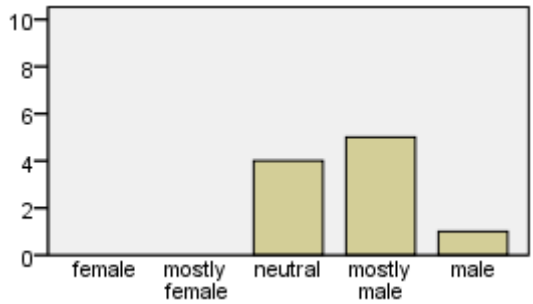


Figure B-4.69. Absolute number of ratings in each category for 'Hung'

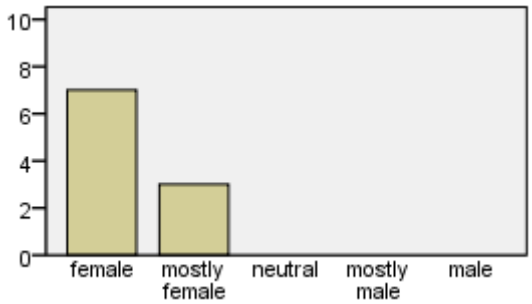


Figure B-4.70. Absolute number of ratings in each category for 'Ida'

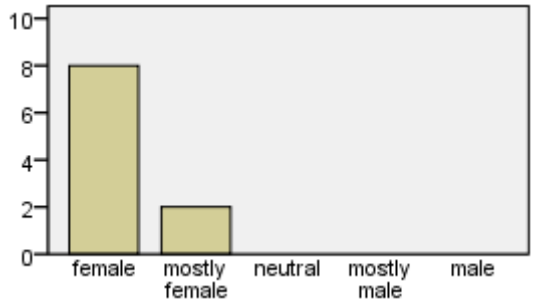


Figure B-4.71. Absolute number of ratings in each category for 'Ingrid'

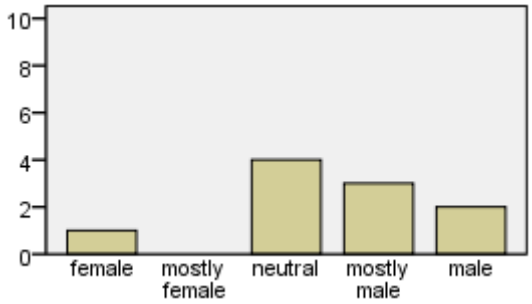


Figure B-4.72. Absolute number of ratings in each category for 'Irfan'

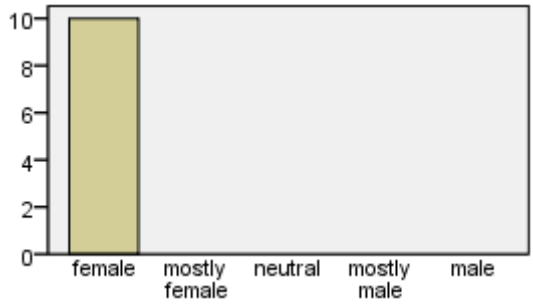


Figure B-4.73. Absolute number of ratings in each category for 'Isabell'

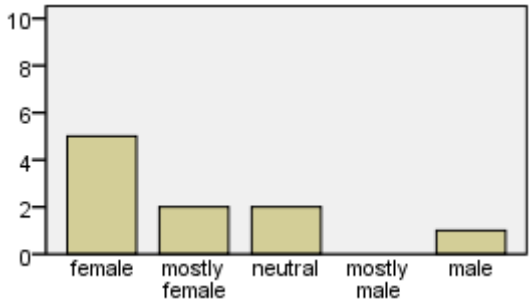


Figure B-4.74. Absolute number of ratings in each category for 'Jade'

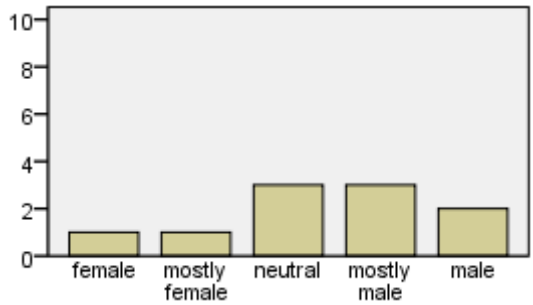


Figure B-4.75. Absolute number of ratings in each category for 'Jaime'

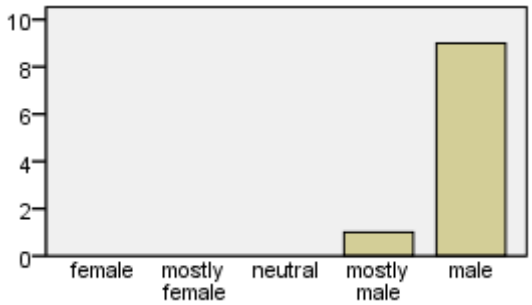


Figure B-4.76. Absolute number of ratings in each category for 'Jan'

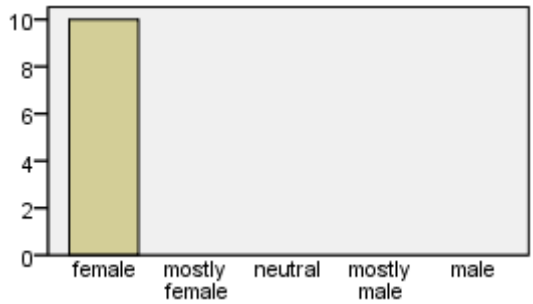


Figure B-4.77. Absolute number of ratings in each category for 'Jana'

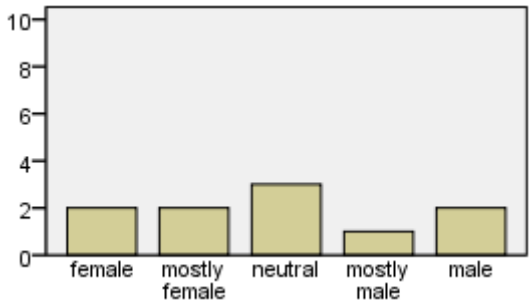


Figure B-4.78. Absolute number of ratings in each category for 'Janis'

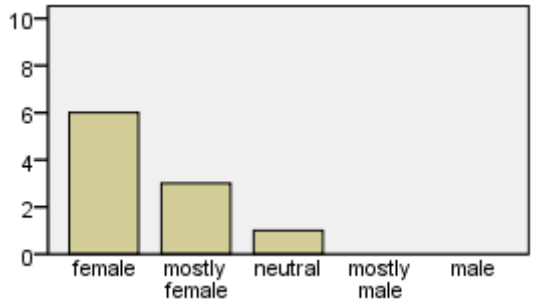


Figure B-4.79. Absolute number of ratings in each category for 'Jemina'

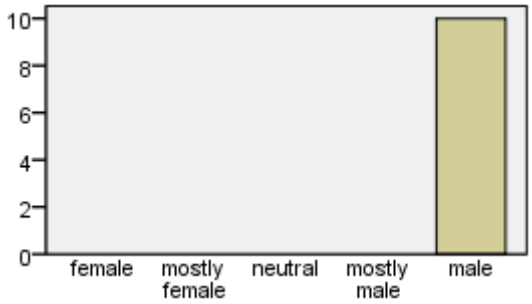


Figure B-4.80. Absolute number of ratings in each category for 'Johannes'

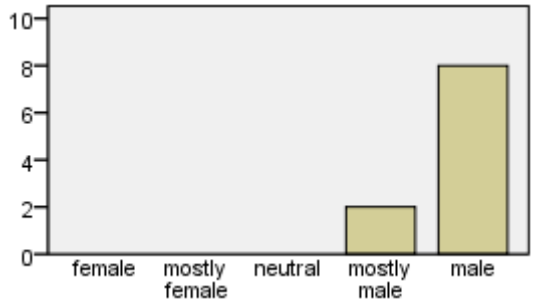


Figure B-4.81. Absolute number of ratings in each category for 'Jonas'

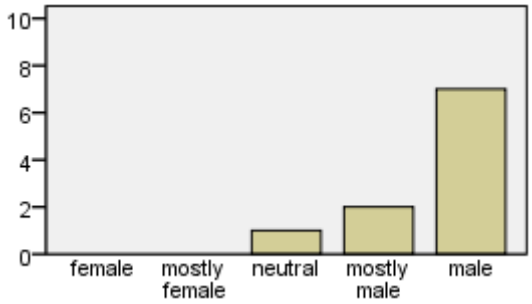


Figure B-4.82. Absolute number of ratings in each category for 'José'

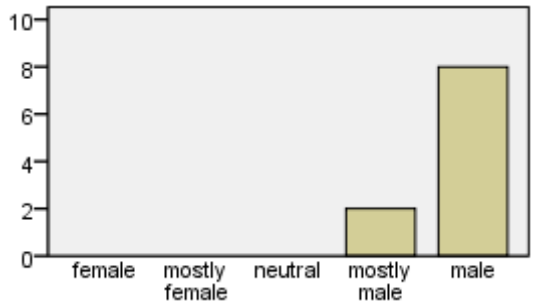


Figure B-4.83. Absolute number of ratings in each category for 'Josua'

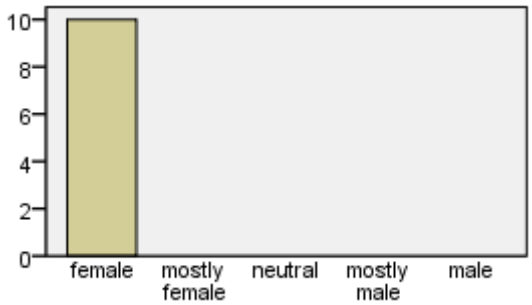


Figure B-4.84. Absolute number of ratings in each category for 'Julia'

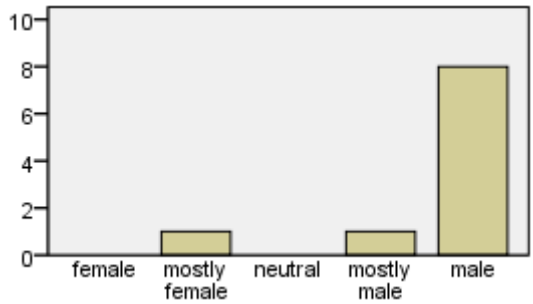


Figure B-4.85. Absolute number of ratings in each category for 'Julian'

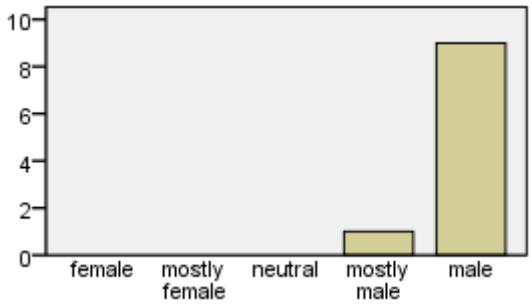


Figure B-4.86. Absolute number of ratings in each category for 'Justin'

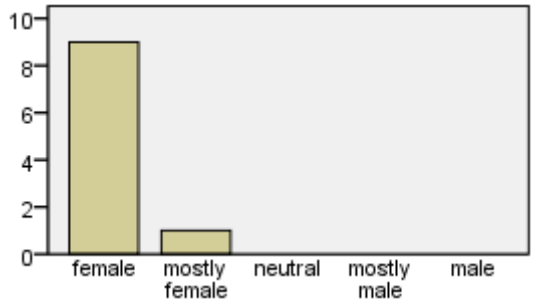


Figure B-4.87. Absolute number of ratings in each category for 'Katharina'

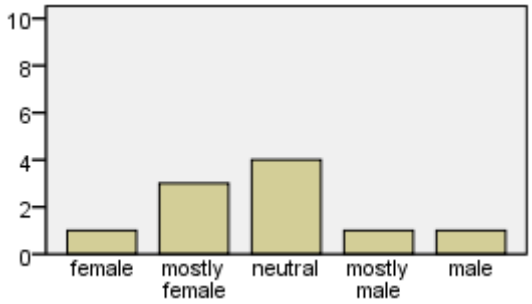


Figure B-4.88. Absolute number of ratings in each category for 'Keiki'

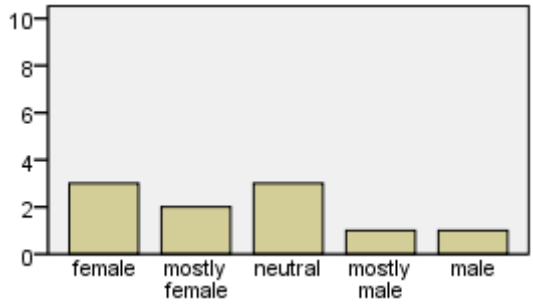


Figure B-4.89. Absolute number of ratings in each category for 'Kim'

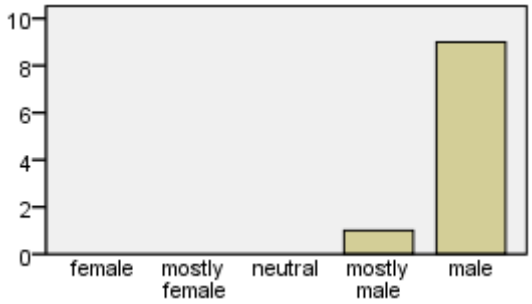


Figure B-4.90. Absolute number of ratings in each category for 'Klemens'

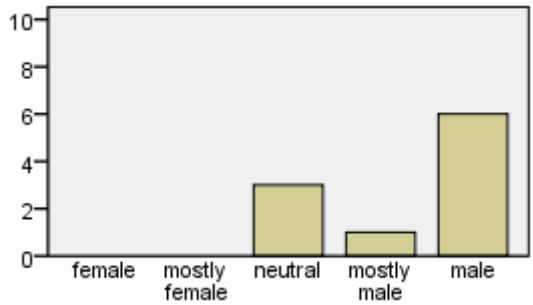


Figure B-4.91. Absolute number of ratings in each category for 'Kyle'

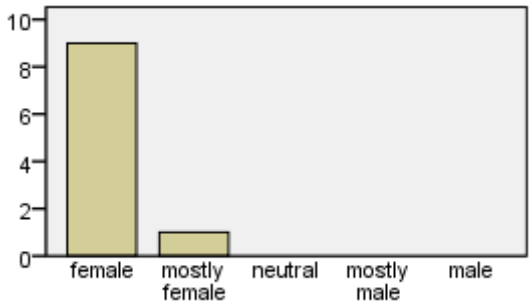


Figure B-4.92. Absolute number of ratings in each category for 'Lana'

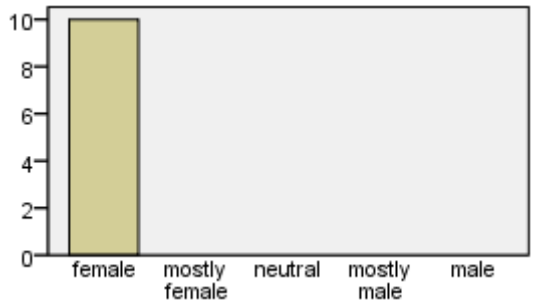


Figure B-4.93. Absolute number of ratings in each category for 'Lara'

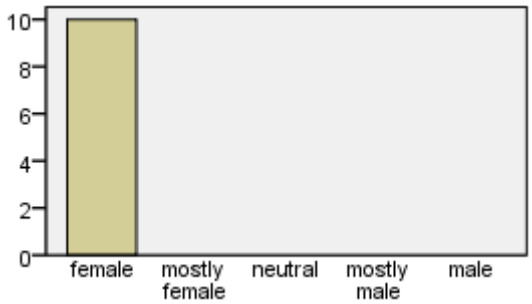


Figure B-4.94. Absolute number of ratings in each category for 'Larissa'

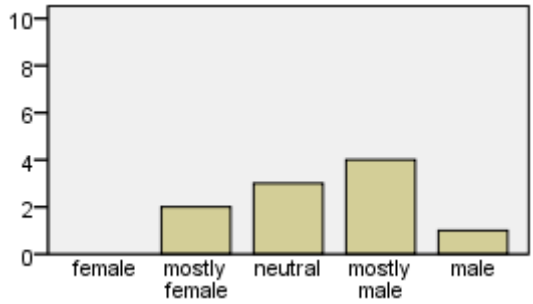


Figure B-4.95. Absolute number of ratings in each category for 'Lærke'

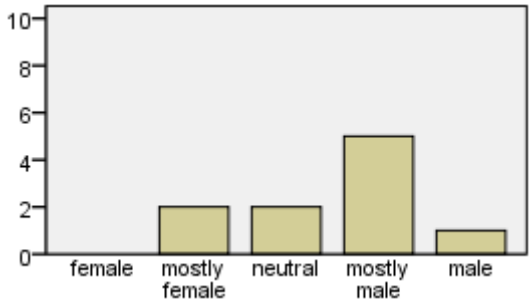


Figure B-4.96. Absolute number of ratings in each category for 'Latif'

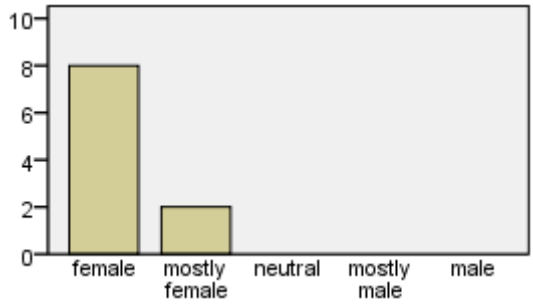


Figure B-4.97. Absolute number of ratings in each category for 'Laura'

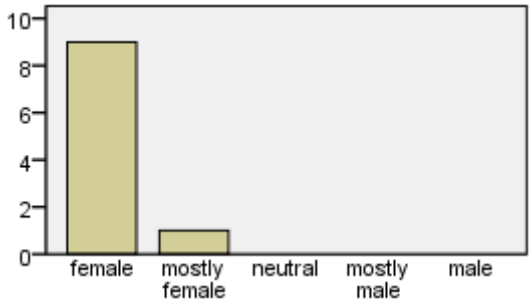


Figure B-4.98. Absolute number of ratings in each category for 'Lena'

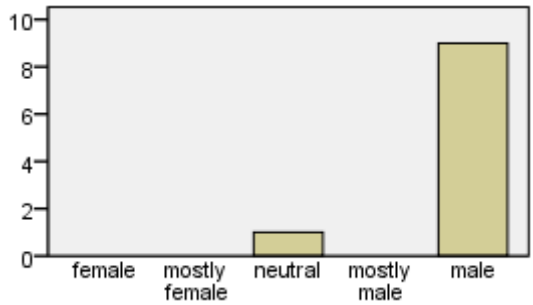


Figure B-4.99. Absolute number of ratings in each category for 'Leon'

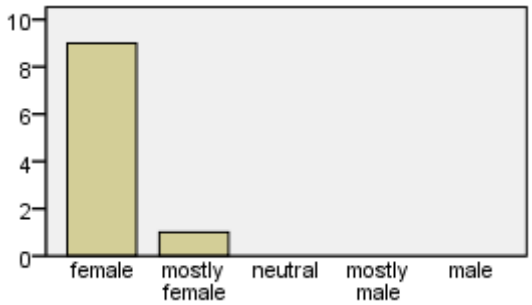


Figure B-4.100. Absolute number of ratings in each category for 'Leonie'

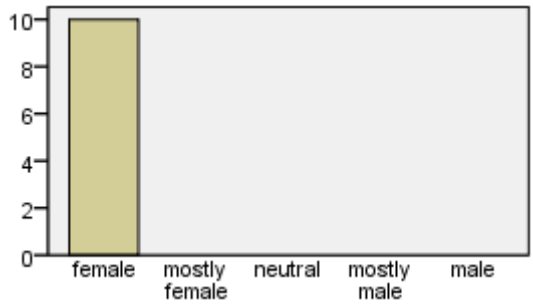


Figure B-4.101. Absolute number of ratings in each category for 'Leyla'

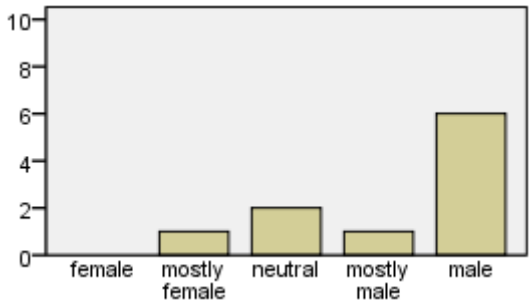


Figure B-4.102. Absolute number of ratings in each category for 'Liam'

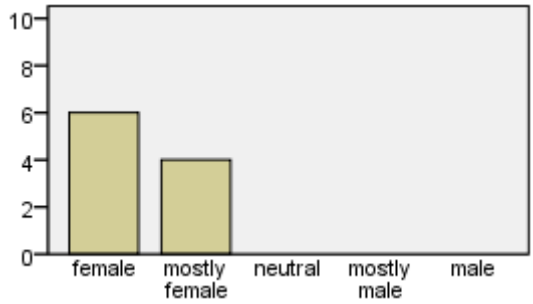


Figure B-4.103. Absolute number of ratings in each category for 'Luoanne'

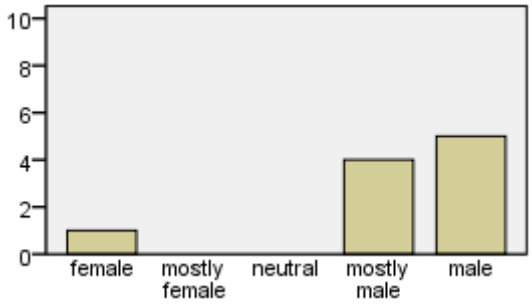


Figure B-4.104. Absolute number of ratings in each category for 'Louis'

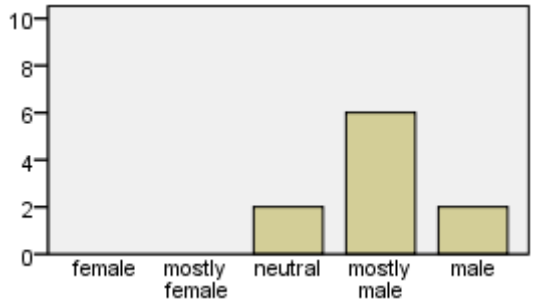


Figure B-4.105. Absolute number of ratings in each category for 'Lovro'

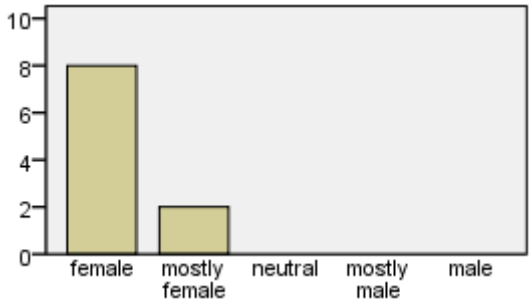


Figure B-4.106. Absolute number of ratings in each category for 'Lucie'

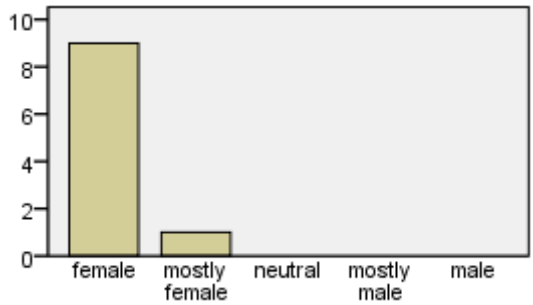


Figure B-4.107. Absolute number of ratings in each category for 'Lucy'

Figure B-4.108. Absolute number of ratings in each category for 'Lukas'

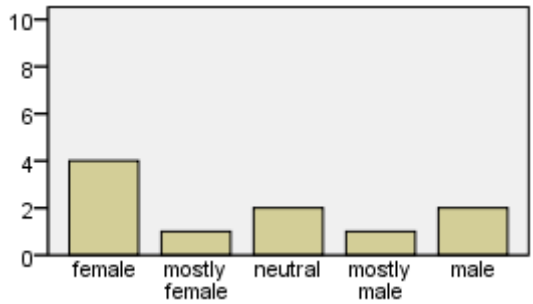


Figure B-4.109. Absolute number of ratings in each category for 'Madison'

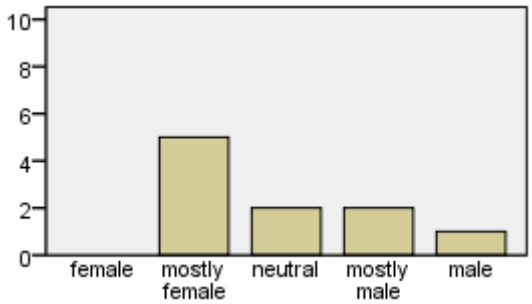


Figure B-4.110. Absolute number of ratings in each category for 'Makani'

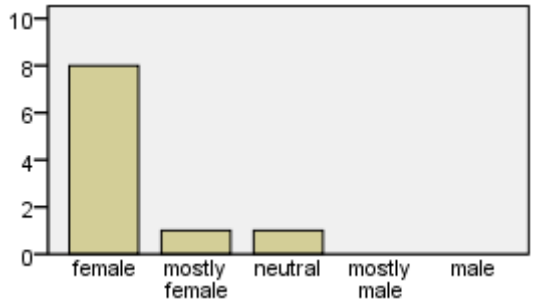


Figure B-4.111. Absolute number of ratings in each category for 'Mara'

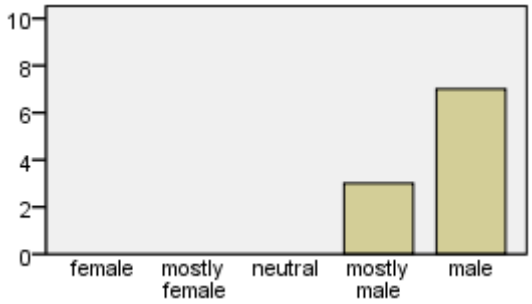


Figure B-4.112. Absolute number of ratings in each category for 'Marcel'

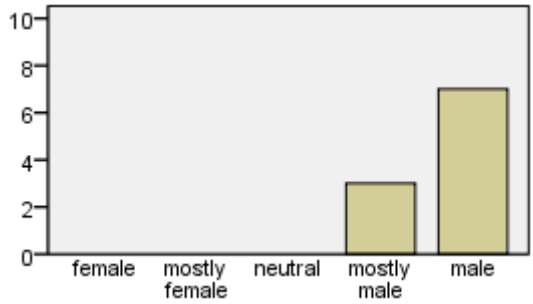


Figure B-4.113. Absolute number of ratings in each category for 'Marek'

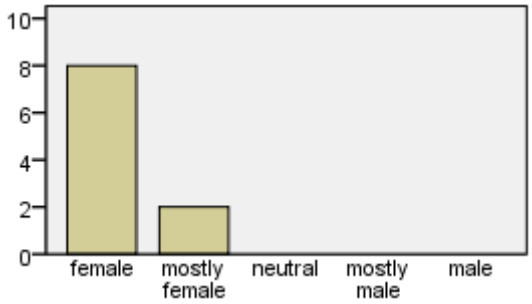


Figure B-4.114. Absolute number of ratings in each category for 'Marie'

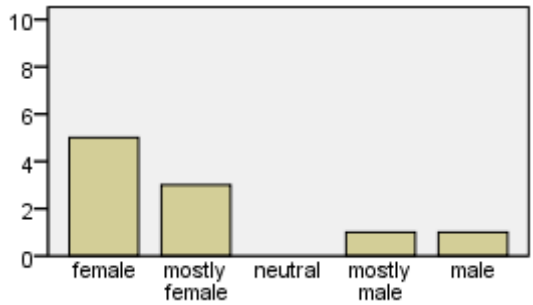


Figure B-4.115. Absolute number of ratings in each category for 'Marit'

Figure B-4.116. Absolute number of ratings in each category for 'Markus'

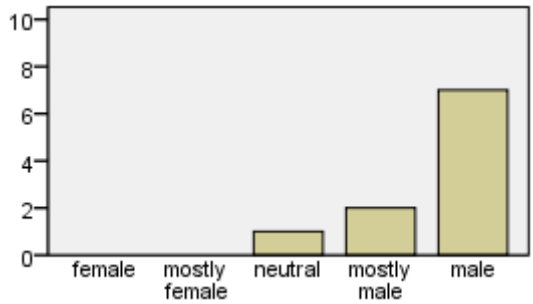


Figure B-4.117. Absolute number of ratings in each category for 'Matej'

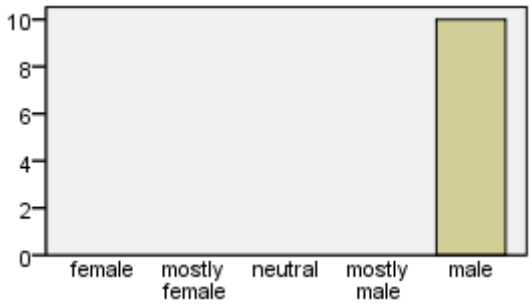


Figure B-4.118. Absolute number of ratings in each category for 'Matthias'

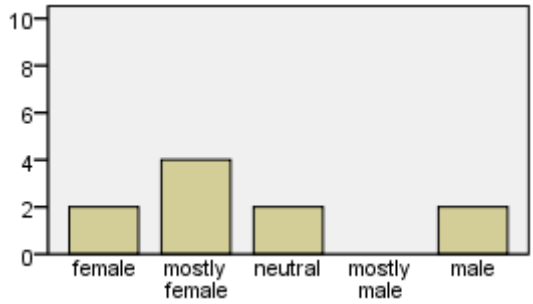


Figure B-4.119. Absolute number of ratings in each category for 'Maxime'

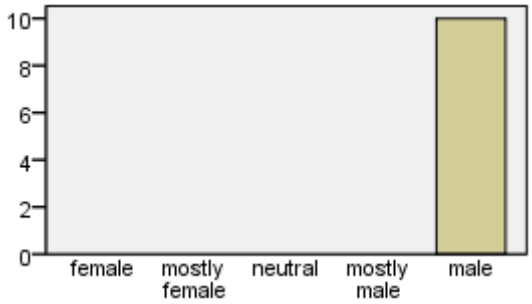


Figure B-4.120. Absolute number of ratings in each category for 'Maximilian'

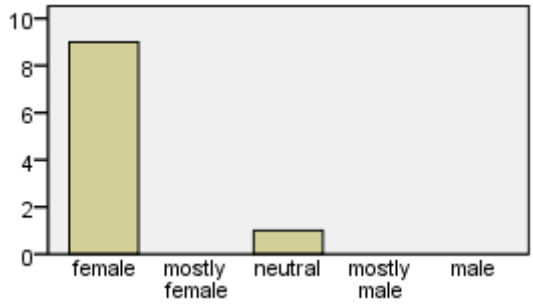


Figure B-4.121. Absolute number of ratings in each category for 'Maya'

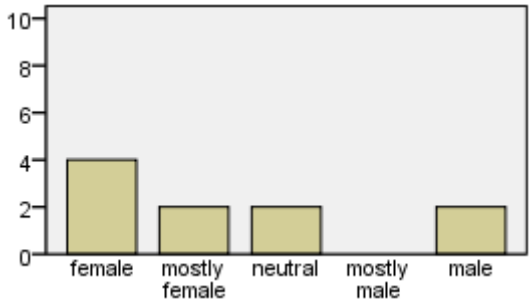


Figure B-4.122. Absolute number of ratings in each category for 'Mercedes'

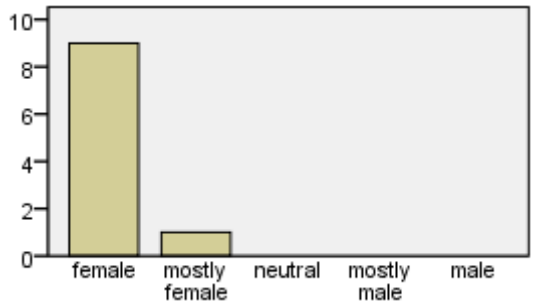


Figure B-4.123. Absolute number of ratings in each category for 'Mia'

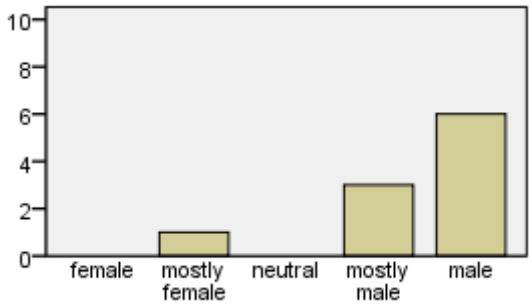


Figure B-4.124. Absolute number of ratings in each category for 'Michal'

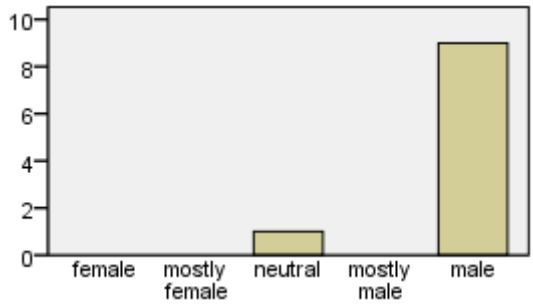


Figure B-4.125. Absolute number of ratings in each category for 'Mikkel'

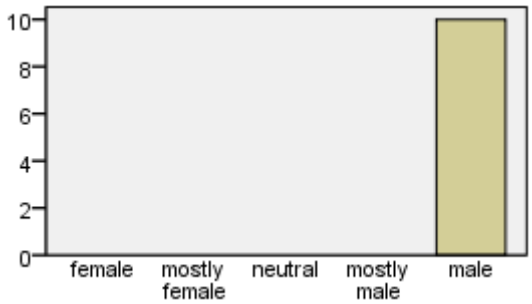


Figure B-4.126. Absolute number of ratings in each category for 'Moritz'

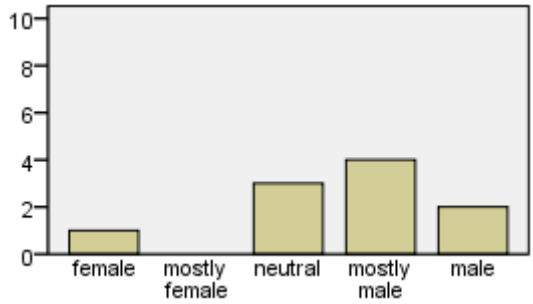


Figure B-4.127. Absolute number of ratings in each category for 'Nadim'

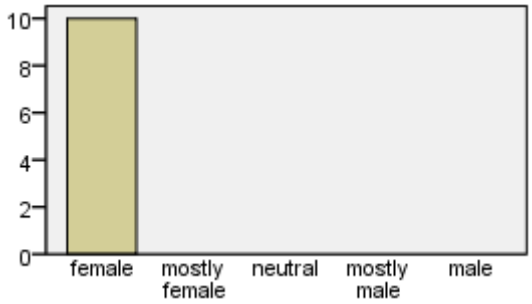


Figure B-4.128. Absolute number of ratings in each category for 'Nadine'

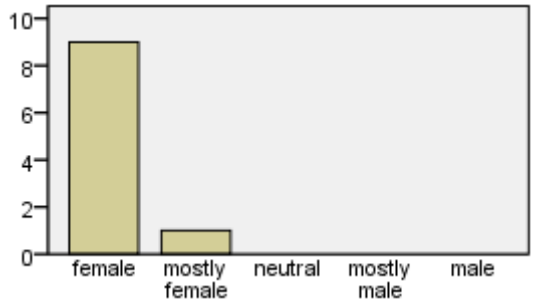


Figure B-4.129. Absolute number of ratings in each category for 'Natalie'

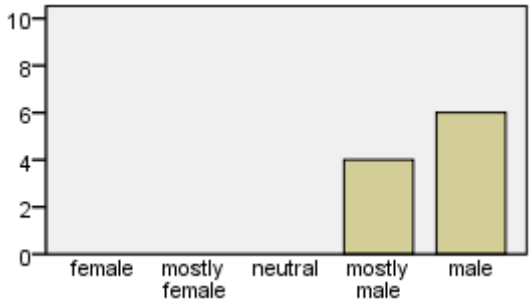


Figure B-4.130. Absolute number of ratings in each category for 'Nathan'

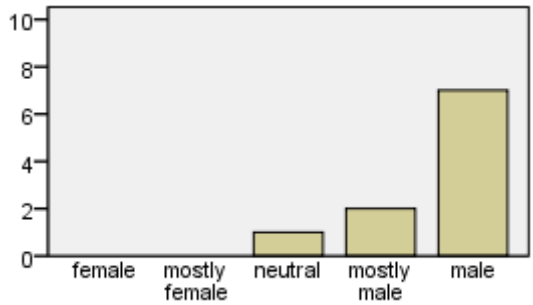


Figure B-4.131. Absolute number of ratings in each category for 'Nico'

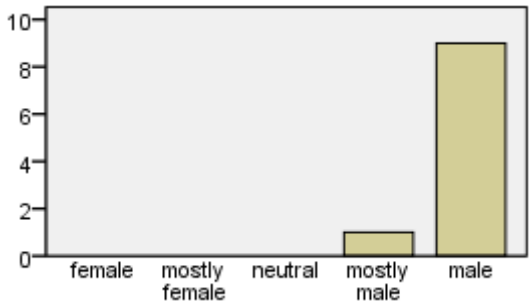


Figure B-4.132. Absolute number of ratings in each category for 'Nikolaj'

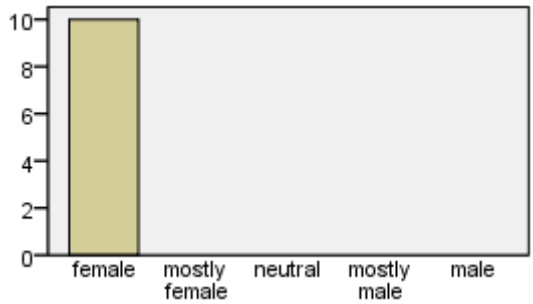


Figure B-4.133. Absolute number of ratings in each category for 'Nina'

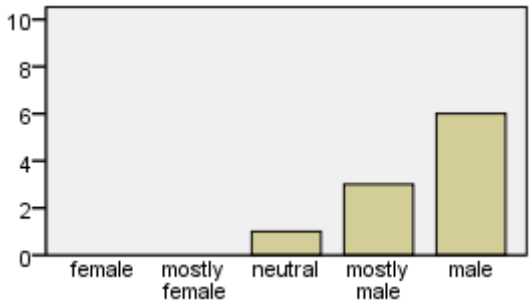


Figure B-4.134. Absolute number of ratings in each category for 'Noah'

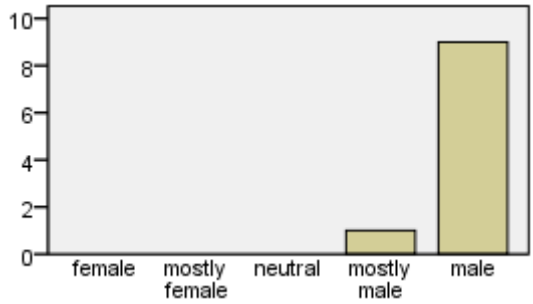


Figure B-4.135. Absolute number of ratings in each category for 'Oliver'

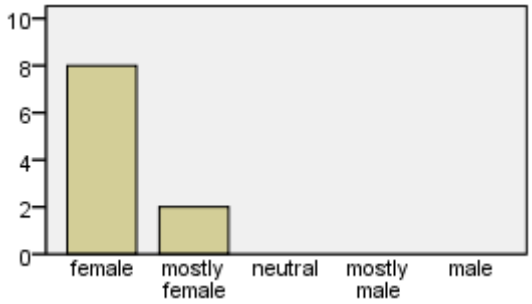


Figure B-4.136. Absolute number of ratings in each category for 'Olivia'

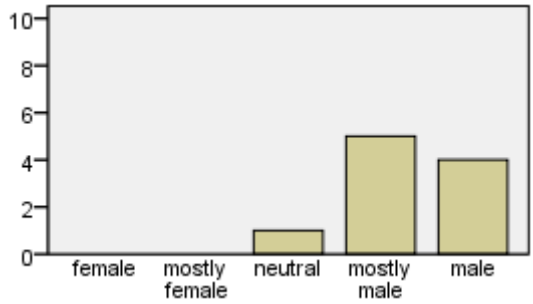


Figure B-4.137. Absolute number of ratings in each category for 'Olle'

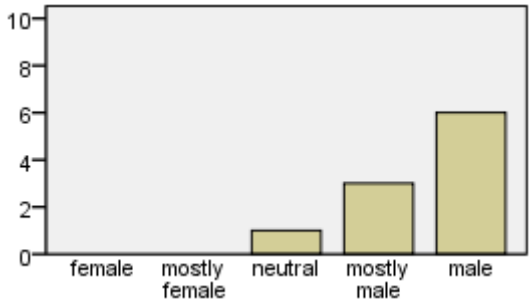


Figure B-4.138. Absolute number of ratings in each category for 'Ondrej'

Figure B-4.139. Absolute number of ratings in each category for 'Osman'

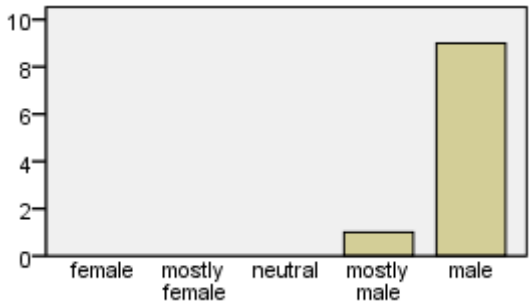


Figure B-4.140. Absolute number of ratings in each category for 'Otto'

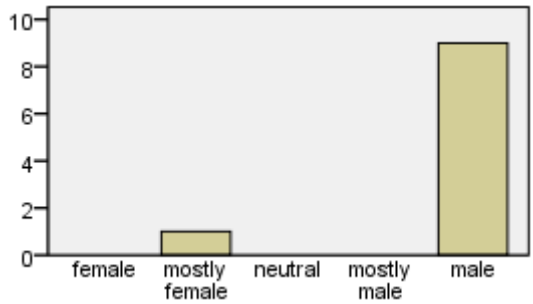


Figure B-4.141. Absolute number of ratings in each category for 'Patrick'

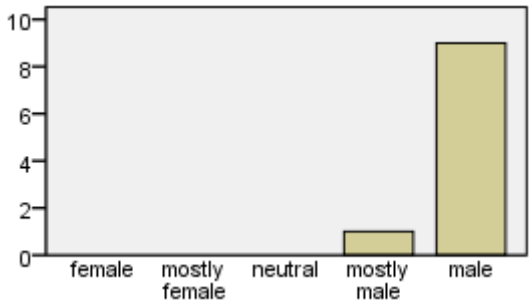


Figure B-4.142. Absolute number of ratings in each category for 'Pedro'

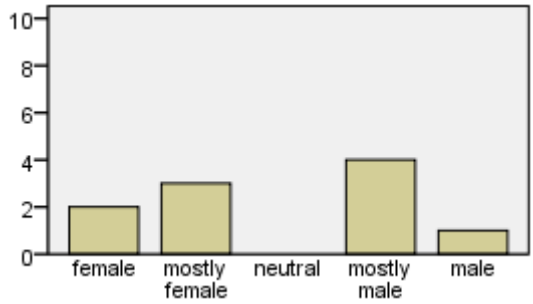


Figure B-4.143. Absolute number of ratings in each category for 'Pinar'

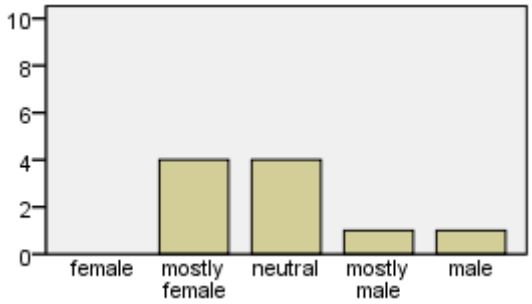


Figure B-4.144. Absolute number of ratings in each category for 'Pua'

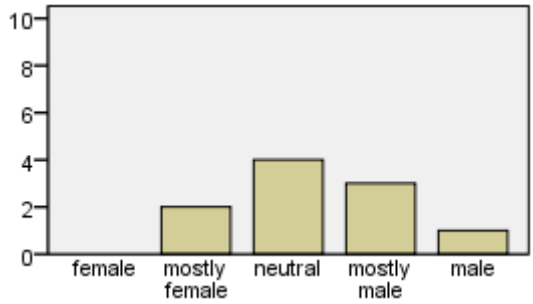


Figure B-4.145. Absolute number of ratings in each category for 'Pualani'

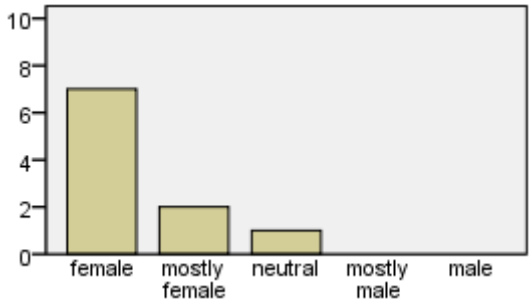


Figure B-4.146. Absolute number of ratings in each category for 'Rachel'

Figure B-4.147. Absolute number of ratings in each category for 'Richard'

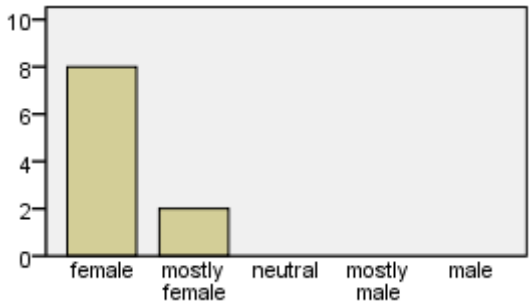


Figure B-4.148. Absolute number of ratings in each category for 'Rita'

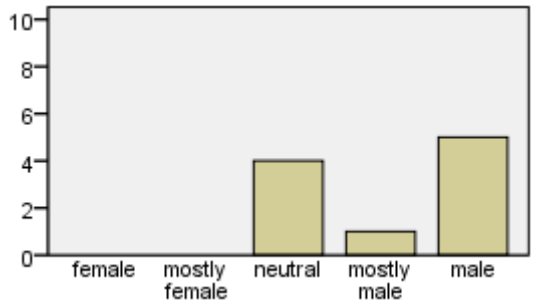


Figure B-4.149. Absolute number of ratings in each category for 'Robin'

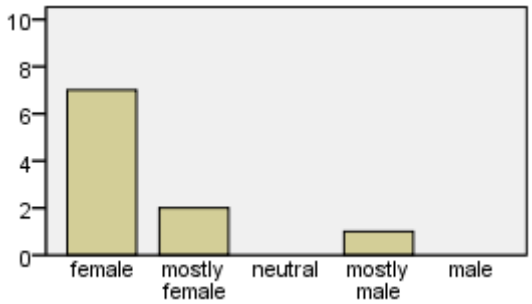


Figure B-4.150. Absolute number of ratings in each category for 'Ronja'

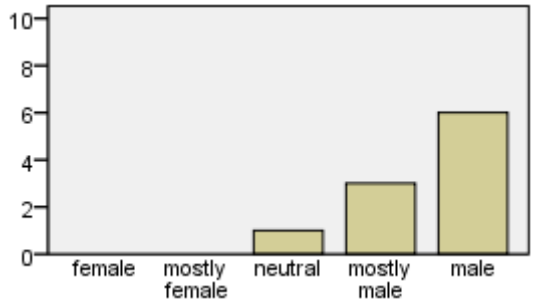


Figure B-4.151. Absolute number of ratings in each category for 'Roque'

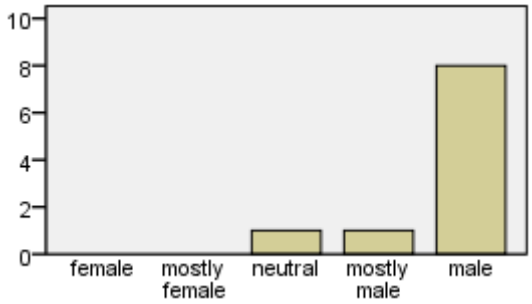


Figure B-4.152. Absolute number of ratings in each category for 'Ryan'

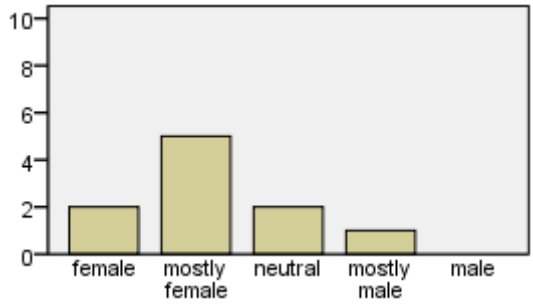


Figure B-4.153. Absolute number of ratings in each category for 'Sabri'

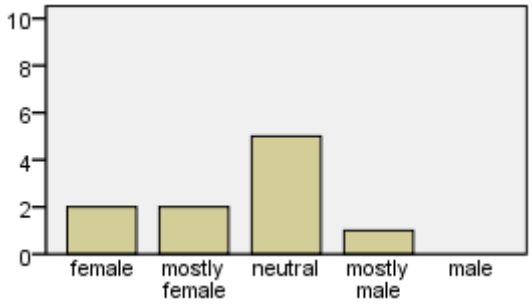


Figure B-4.154. Absolute number of ratings in each category for 'Safa'

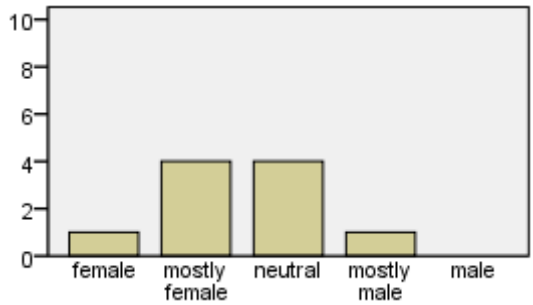


Figure B-4.155. Absolute number of ratings in each category for 'Saga'

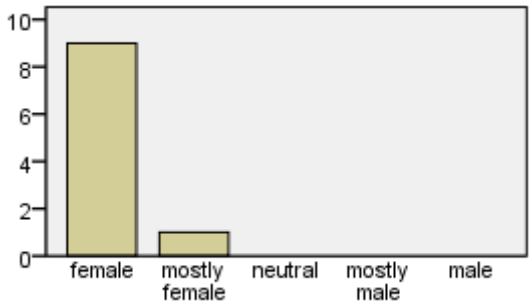


Figure B-4.156. Absolute number of ratings in each category for 'Sarah'

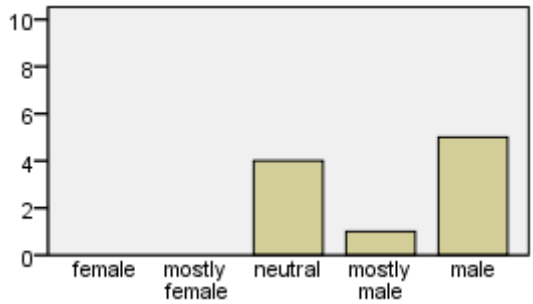


Figure B-4.157. Absolute number of ratings in each category for 'Sascha'

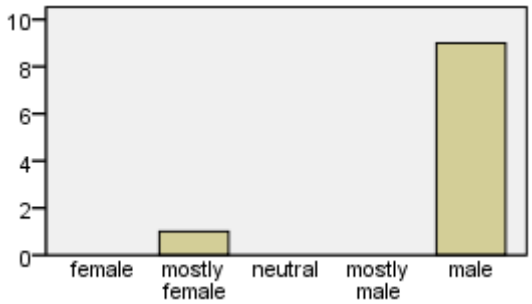


Figure B-4.158. Absolute number of ratings in each category for 'Sebastian'

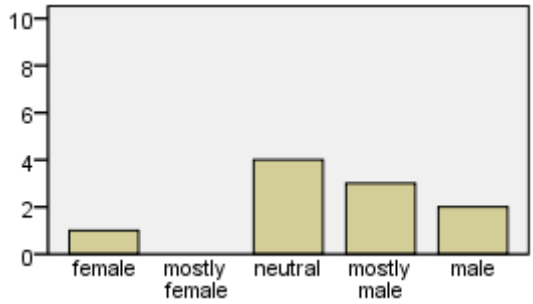


Figure B-4.159. Absolute number of ratings in each category for 'Seher'

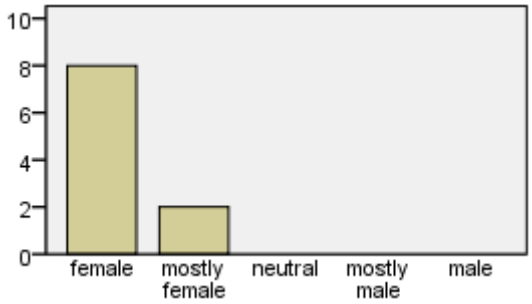


Figure B-4.160. Absolute number of ratings in each category for 'Selina'

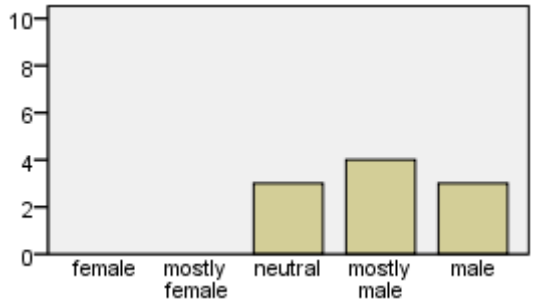


Figure B-4.161. Absolute number of ratings in each category for 'Seval'

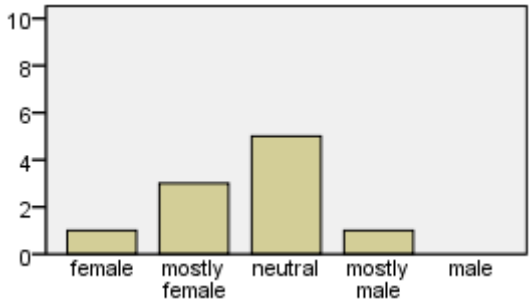


Figure B-4.162. Absolute number of ratings in each category for 'Sidney'

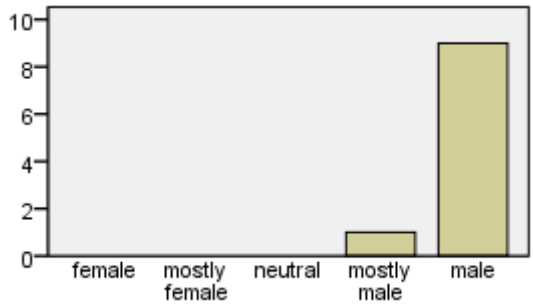


Figure B-4.163. Absolute number of ratings in each category for 'Simon'

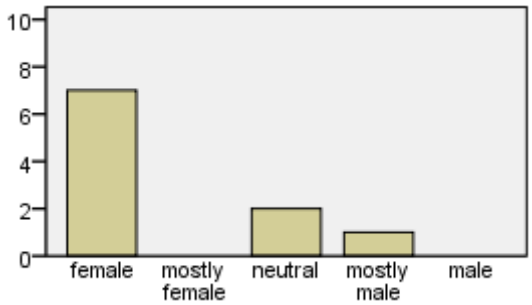


Figure B-4.164. Absolute number of ratings in each category for 'Simone'

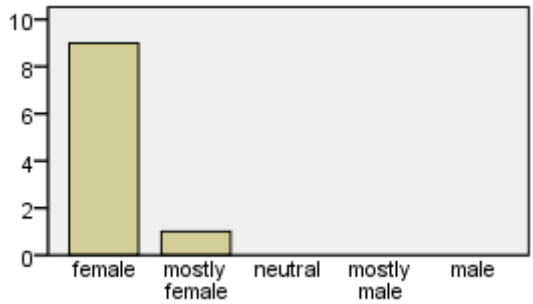


Figure B-4.165. Absolute number of ratings in each category for 'Sofia'

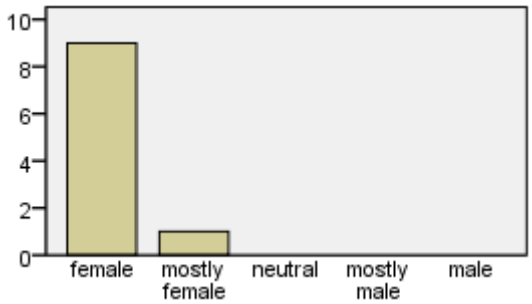


Figure B-4.166. Absolute number of ratings in each category for 'Sophie'

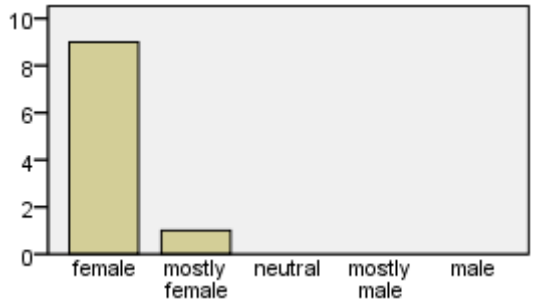


Figure B-4.167. Absolute number of ratings in each category for 'Susanne'

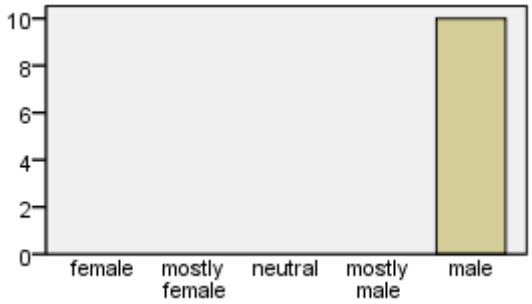


Figure B-4.168. Absolute number of ratings in each category for 'Sven'

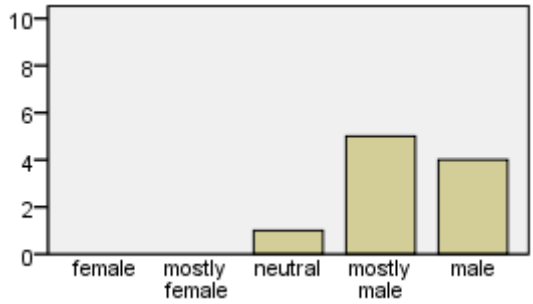


Figure B-4.169. Absolute number of ratings in each category for 'Tahir'

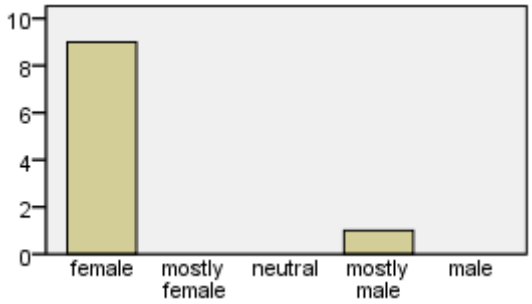


Figure B-4.170. Absolute number of ratings in each category for 'Tanja'

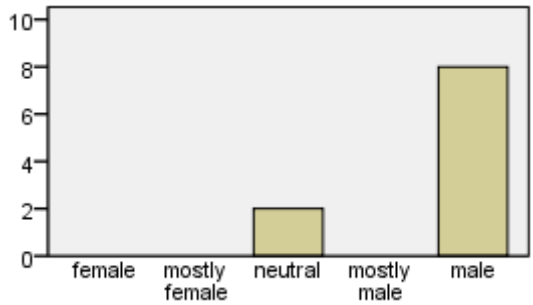


Figure B-4.171. Absolute number of ratings in each category for 'Tarek'

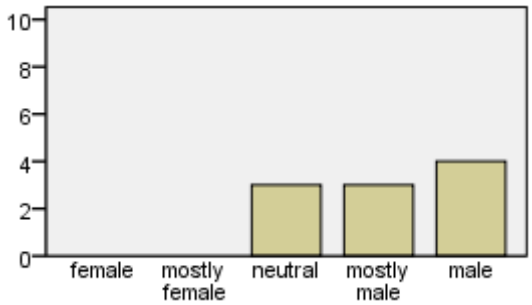


Figure B-4.172. Absolute number of ratings in each category for 'Thao'

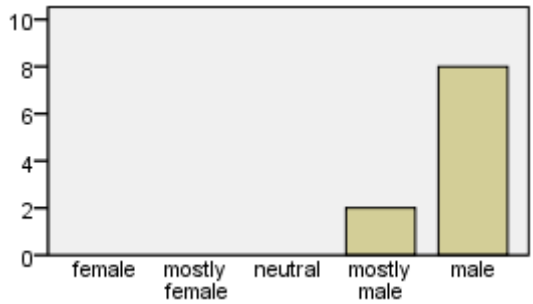


Figure B-4.173. Absolute number of ratings in each category for 'Tim'

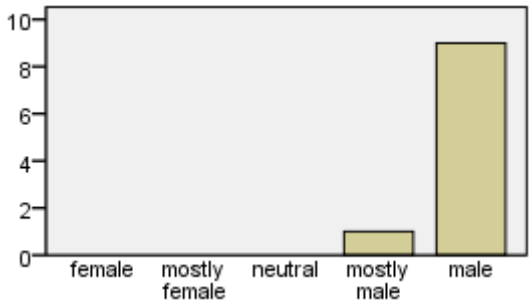


Figure B-4.174. Absolute number of ratings in each category for 'Tobias'

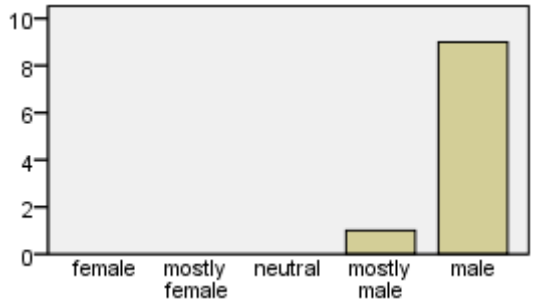


Figure B-4.175. Absolute number of ratings in each category for 'Tom'

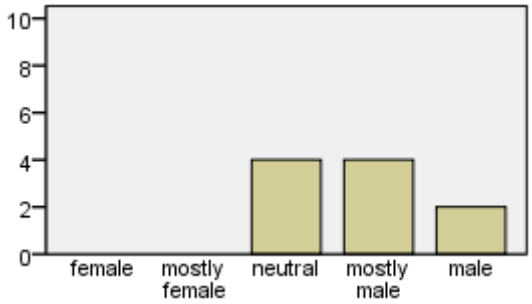


Figure B-4.176. Absolute number of ratings in each category for 'Tung'

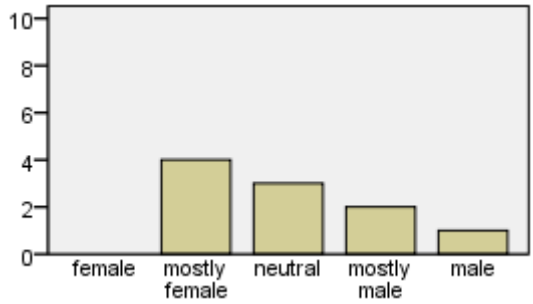


Figure B-4.177. Absolute number of ratings in each category for 'Tuva'

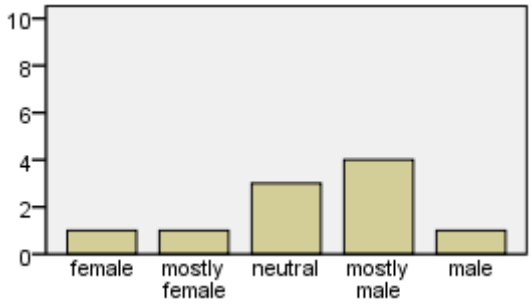


Figure B-4.178. Absolute number of ratings in each category for 'Vega'

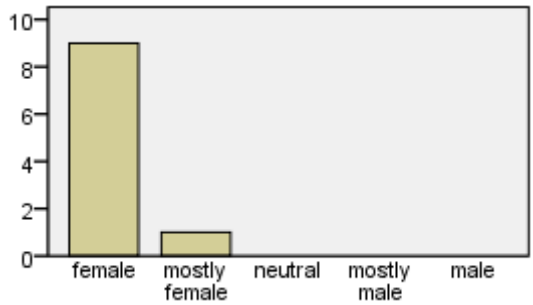


Figure B-4.179. Absolute number of ratings in each category for 'Vera'

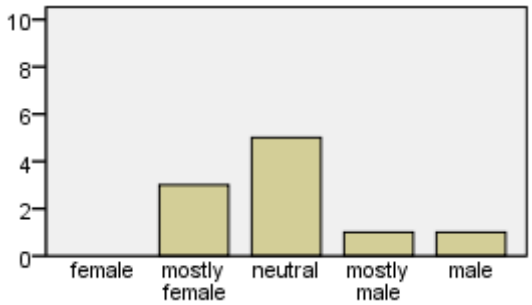


Figure B-4.180. Absolute number of ratings in each category for 'Vilde'

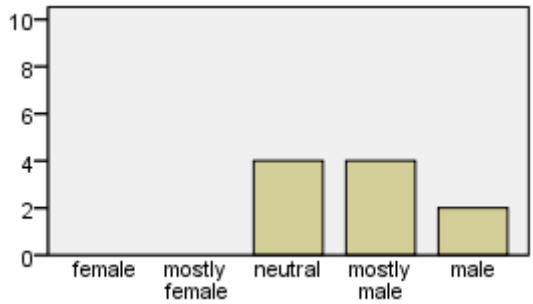


Figure B-4.181. Absolute number of ratings in each category for 'Wedat'

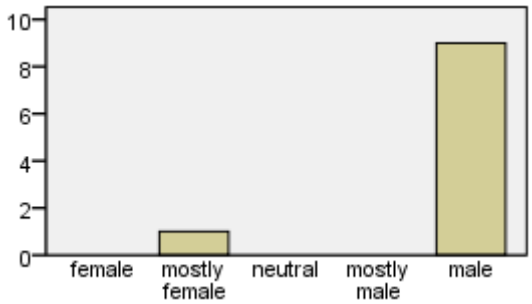


Figure B-4.182. Absolute number of ratings in each category for 'William'

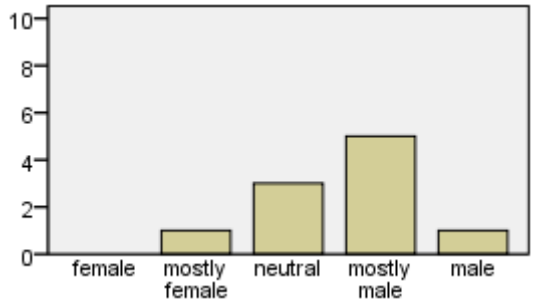


Figure B-4.183. Absolute number of ratings in each category for 'Yael'

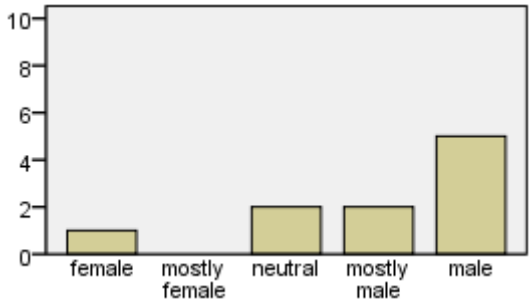


Figure B-4.184. Absolute number of ratings in each category for 'Yanis'

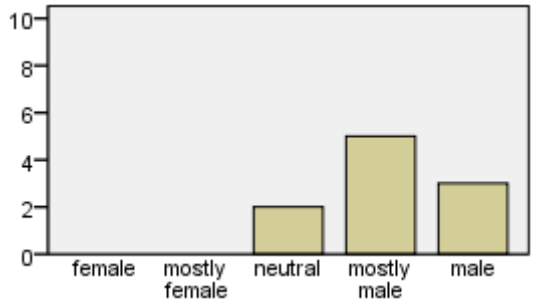


Figure B-4.185. Absolute number of ratings in each category for 'Yasar'

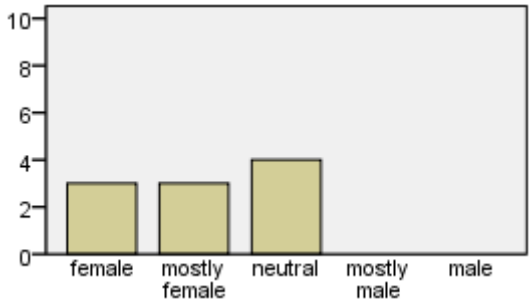


Figure B-4.186. Absolute number of ratings in each category for 'Zada'

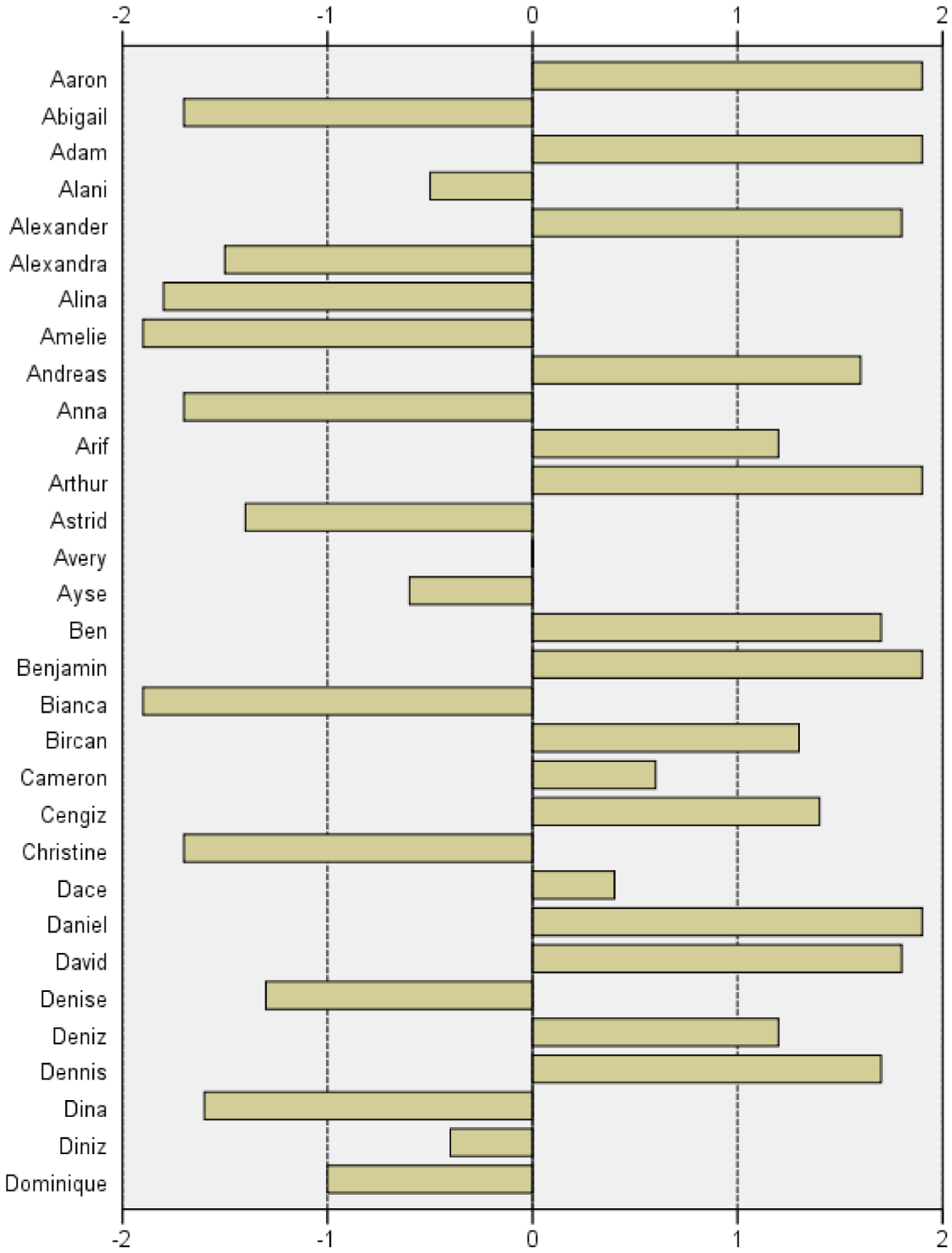


Figure B-4.187a. Bar Chart of Means for the names from Aaron to Dominique (arranged alphabetically).

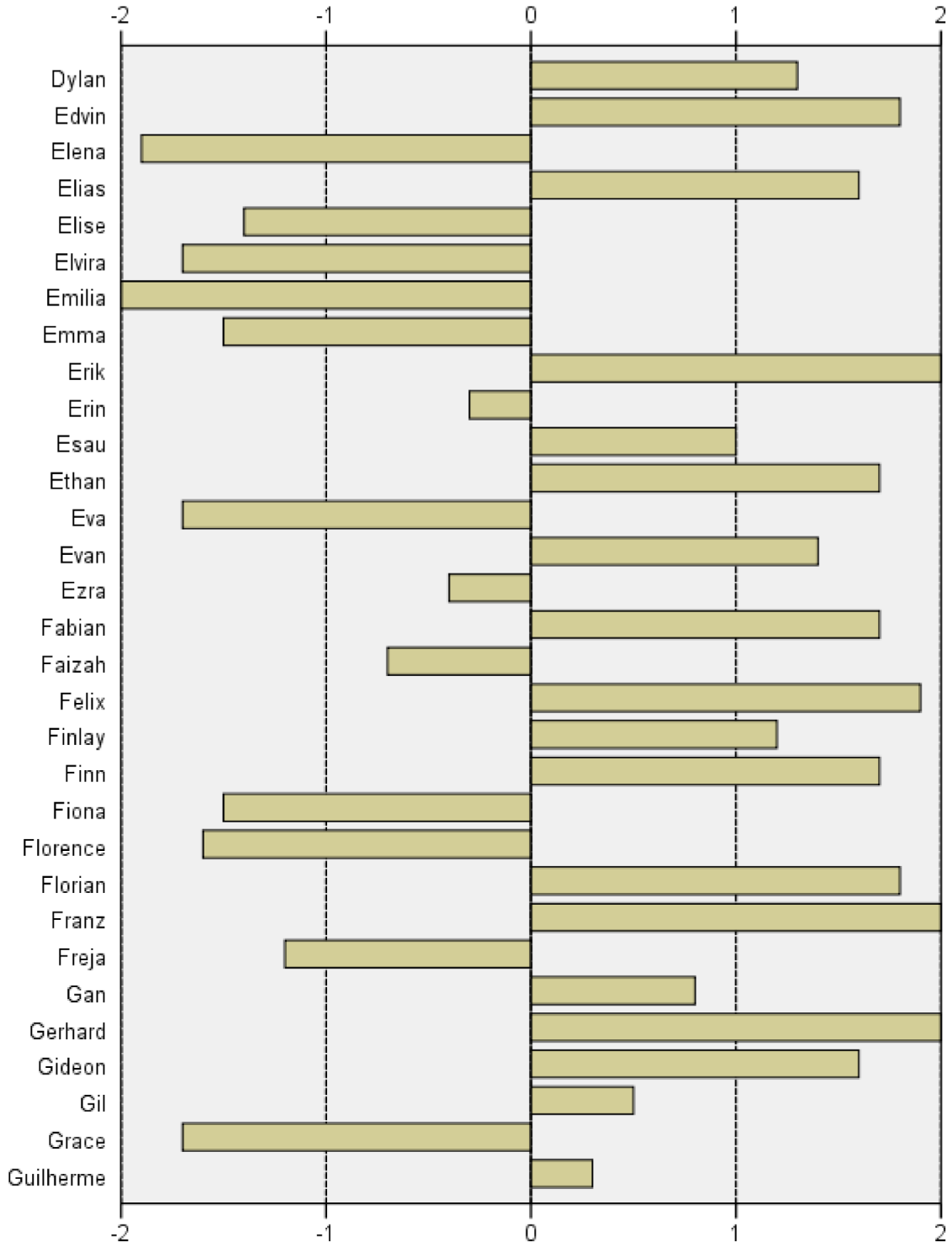


Figure B-4.187b. Bar Chart of Means for the names from Dylan to Guilherme (arranged alphabetically).

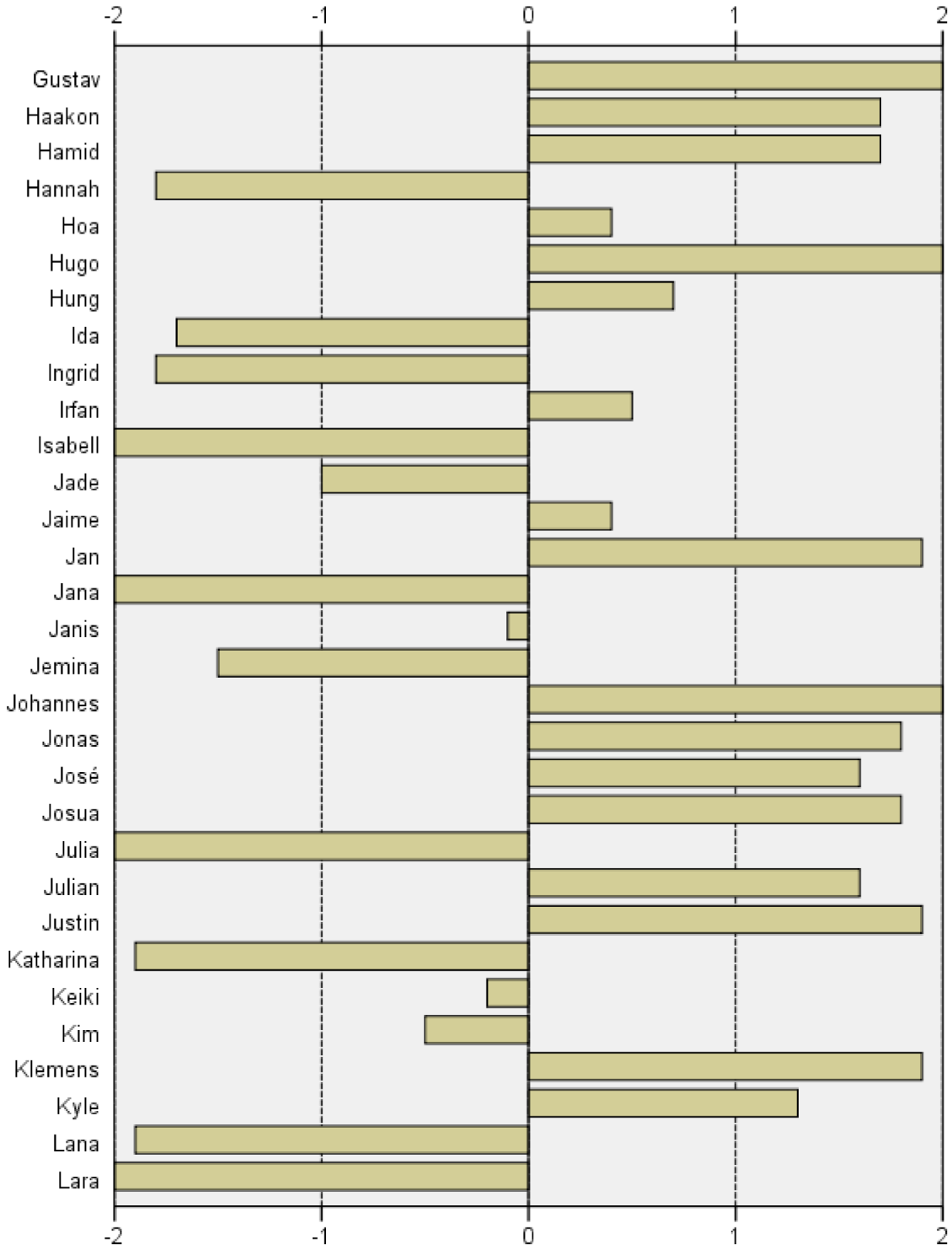


Figure B-4.187c. Bar Chart of Means for the names from Gustav to Lara (arranged alphabetically).

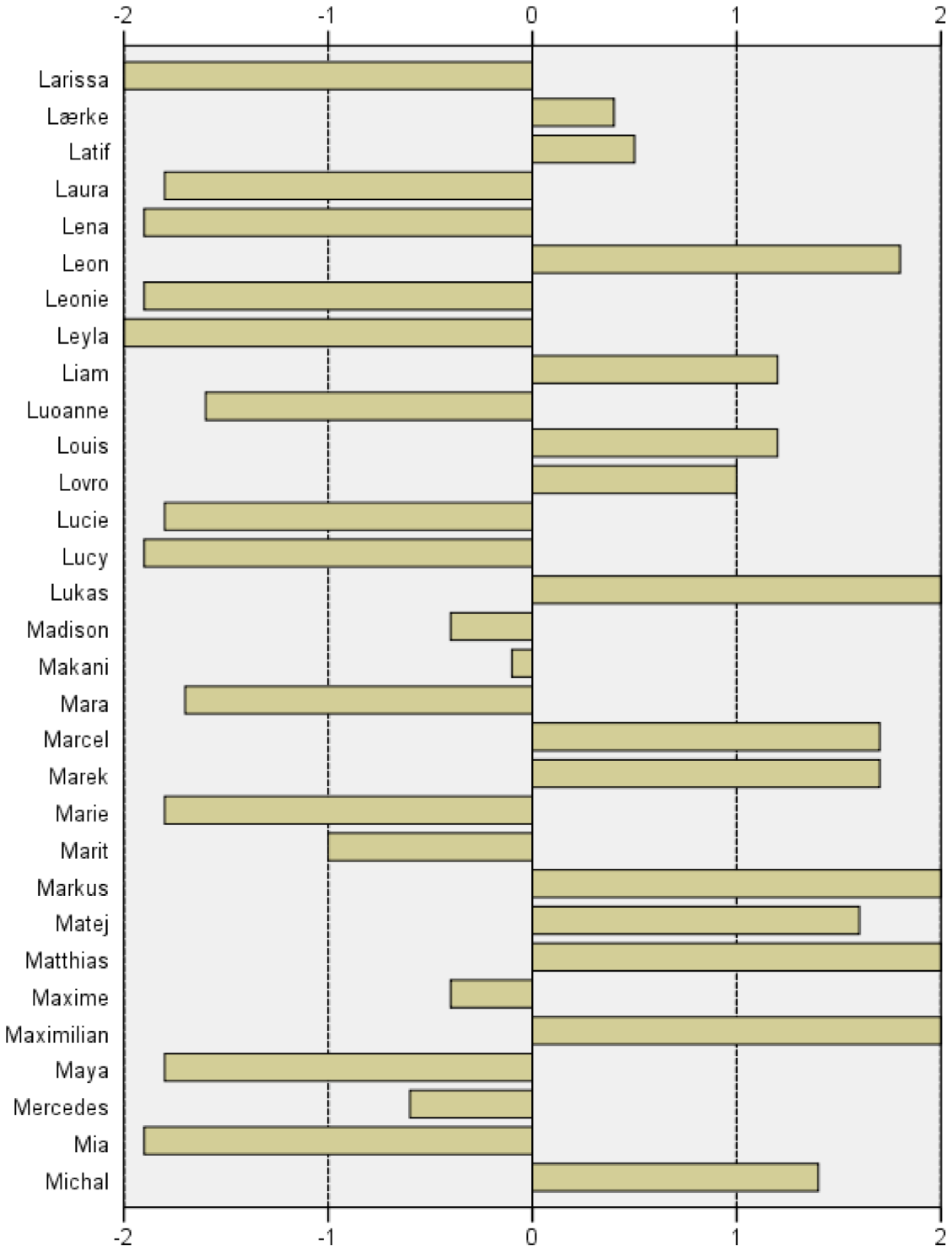


Figure B-4.187d. Bar Chart of Means for the names from Larissa to Michal (arranged alphabetically).

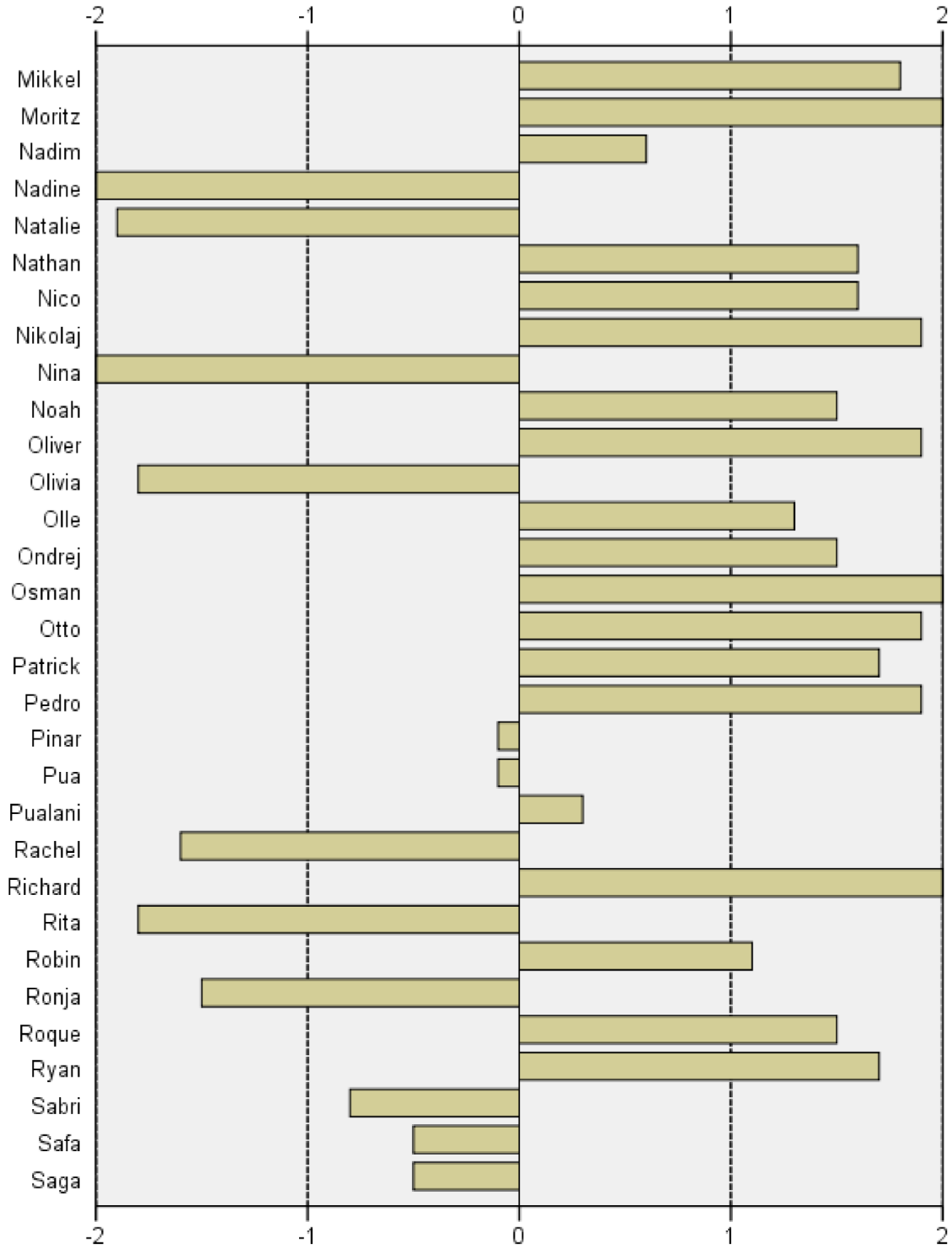


Figure B-4.187e. Bar Chart of Means for the names from Mikkel to Saga (arranged alphabetically).

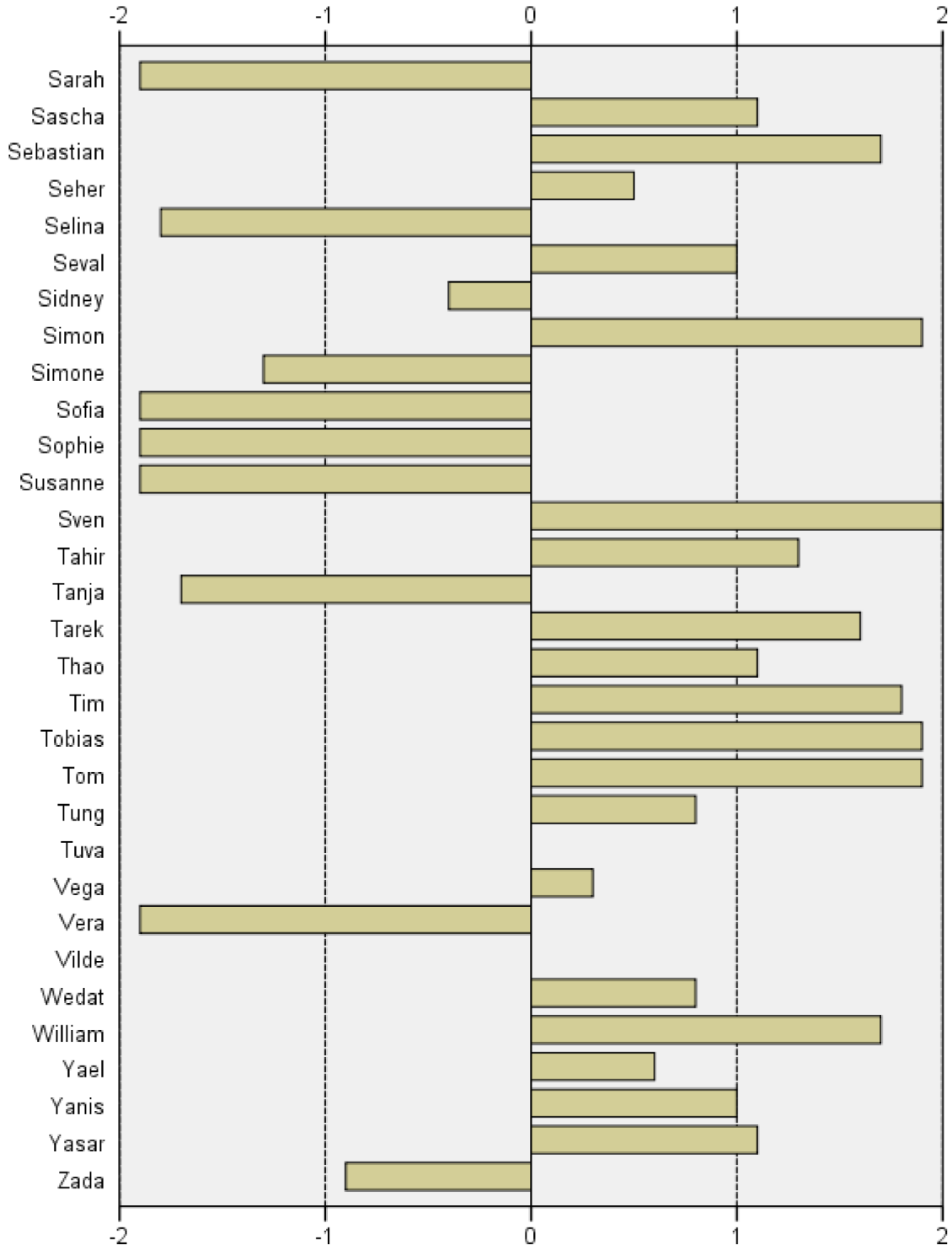


Figure B-4.187f. Bar Chart of Means for the names from Sarah to Zada (arranged alphabetically).

Appendix C0 [Main Experiment – Instructions]

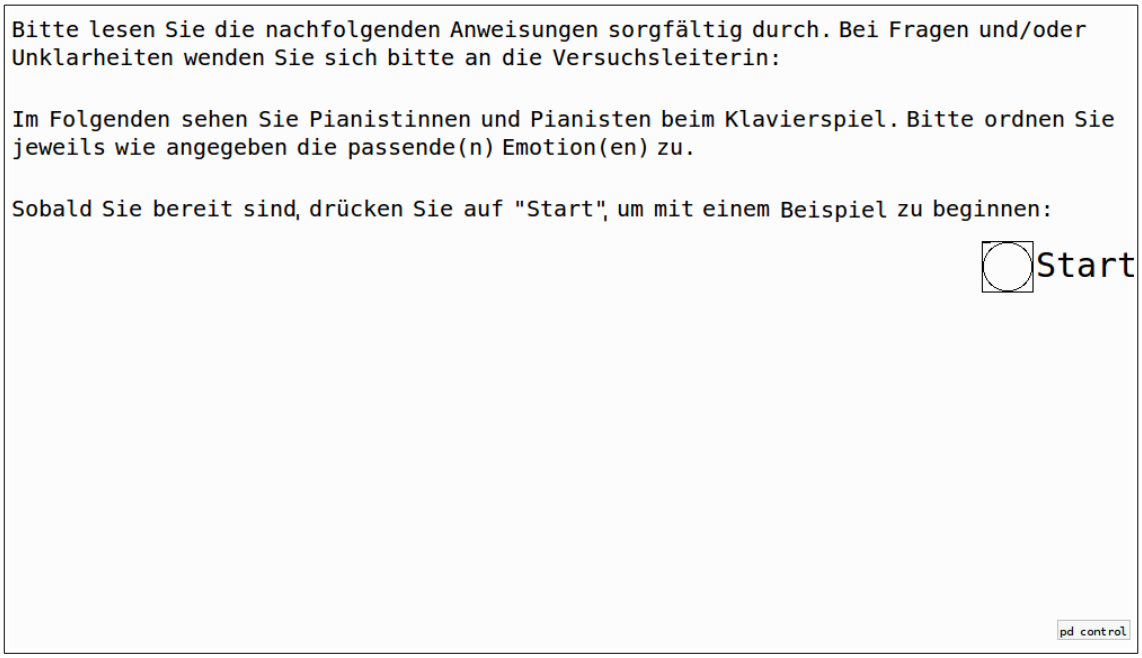


Figure C-0.1. Screenshot of the PureData patch of the main experiment showing the instructions (The female option, 'Versuchsleiterin' is presented in this figure, and dynamically changed into the mail option 'Versuchsleiter' if selected in the control patch).

Own translation of the German instructions:
Please read the following instructions carefully. In case of questions and/or obscurities please ask the investigator.
In the following, you will see female and male pianists while playing. Please assign the suitable emotion(s) to each case, as stated.
Click “Start” as soon as you are ready, to start with an example.

PureData patch available on appendix CD:
[/content/experiments/main_study/main_experiment.pd]

Appendix C1 [Main Experiment – PureData Patch]

The PureData patch of the main experiment is an adjusted and improved version of the patch used in the first pre-experiment. The patch consists of two different groups of sub-patches: Graphical interfaces (Figures C-1. To C-1.) and background/control sub-patches (Figures C-1. To C-1.).

The graphical interfaces served as input masks for the participants. The experiment could be started simply from the participants by clicking on the start button on the instructions sub-patch (Figure C-0.1). All program operation processes were developed using built-in objects of PureData and embedded in the background/control sub-patches: Automatic playing of the videos (after clicking the start/next-button), displaying and hiding different interfaces and parts of input masks, writing the raw data in a text file, and the consecutive numbering. Detailed information about all built-in objects used in this patch can be found in the help file of the software.

In addition to the patch used in the first pre-experiment, this patch has two options to adjust the interface. Both options are located in the control sub-patch (Figure C-1.): It is possible (1) to select the investigator (Johannes/male, or Nadine/female) to adjust the instructions, and (2) to switch between experiment and design mode to adjusted background colours and font for better contrast/readability.

Note. The PureData patch does NOT entirely work without the video files. The videos were exclusively provided for the purpose of this master thesis and, therefore, not included in the appendix or the digital appendix. If the videos are needed, please contact Marc Thompson and ask for permission (if granted, the adjusted video clips can be provided by me).

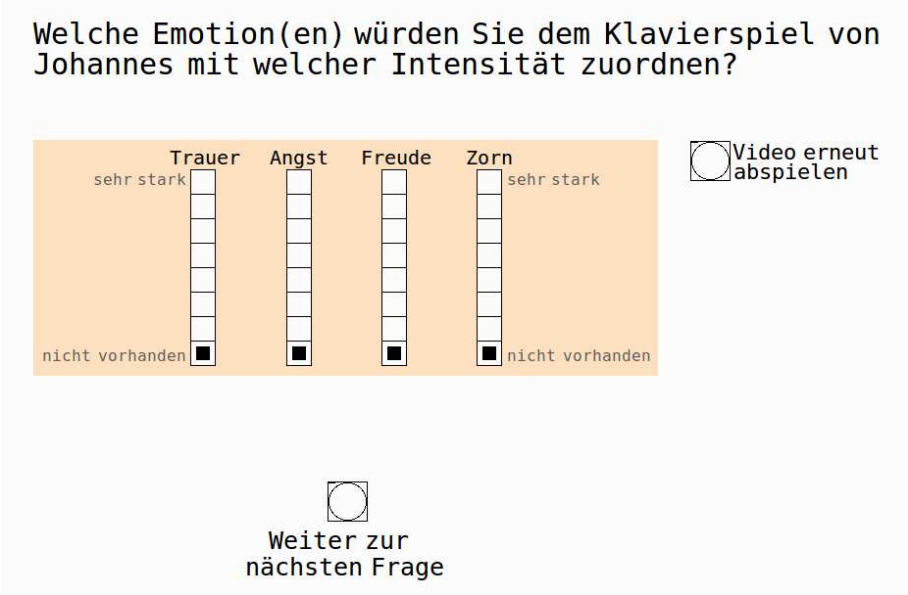


Figure C-1-1. Screenshot showing the input screen of the used PureData patch.

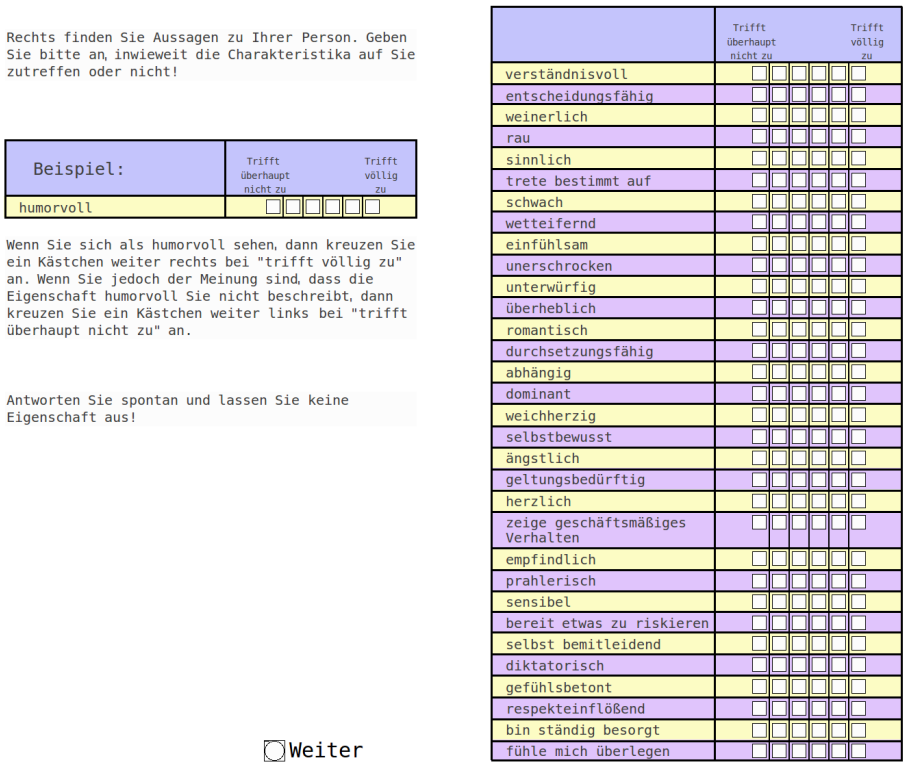


Figure C-1-2. Screenshot showing the German Extended Personal Attributes Questionnaire (GE-PAQ) input screen sub-patch.

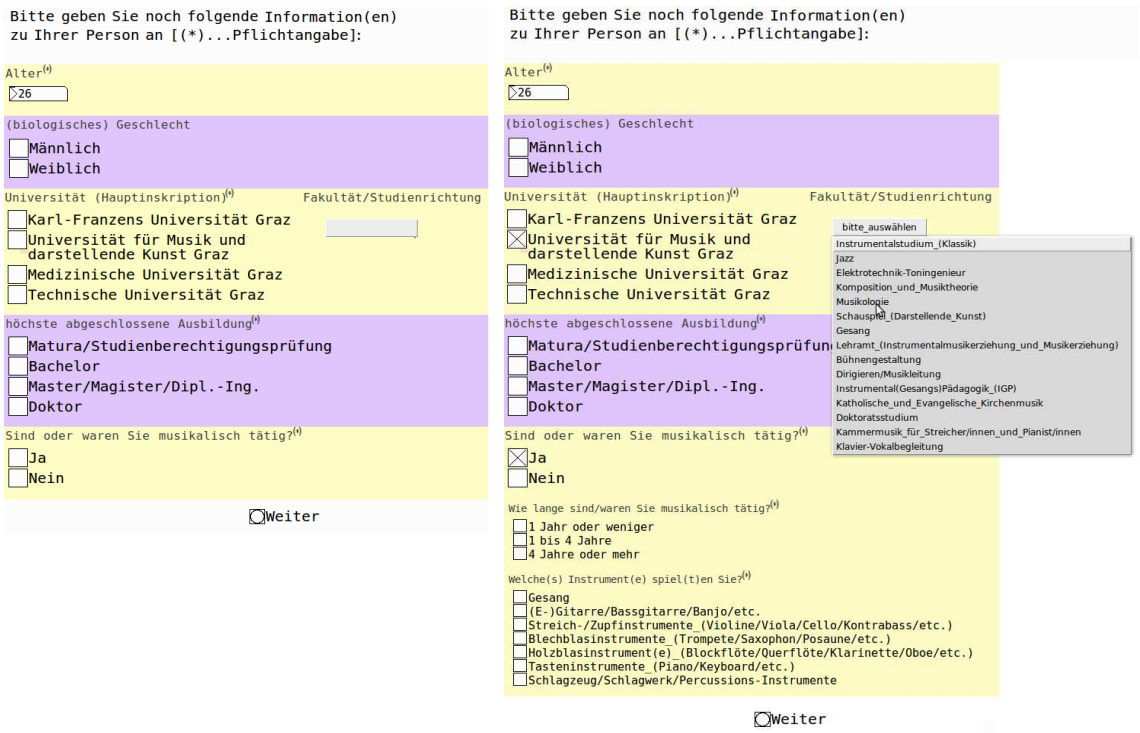


Figure C-1-3. Screenshots showing the demographic input screen. The basic input screen (left), and the extended input screen with additional options to specify the musical education and the faculty/study (right).

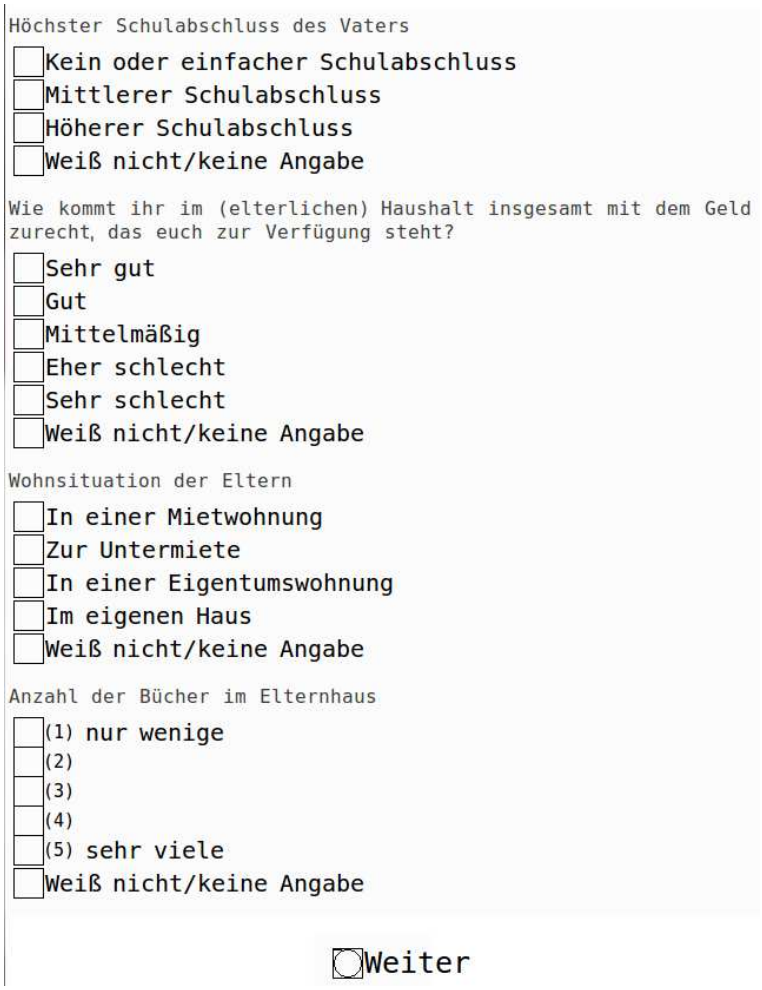


Figure C-1-4. Screenshot showing the social strata input screen sub-patch.

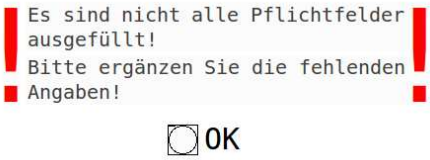


Figure C-1-5. Screenshot showing the alert message of the PureData patch.

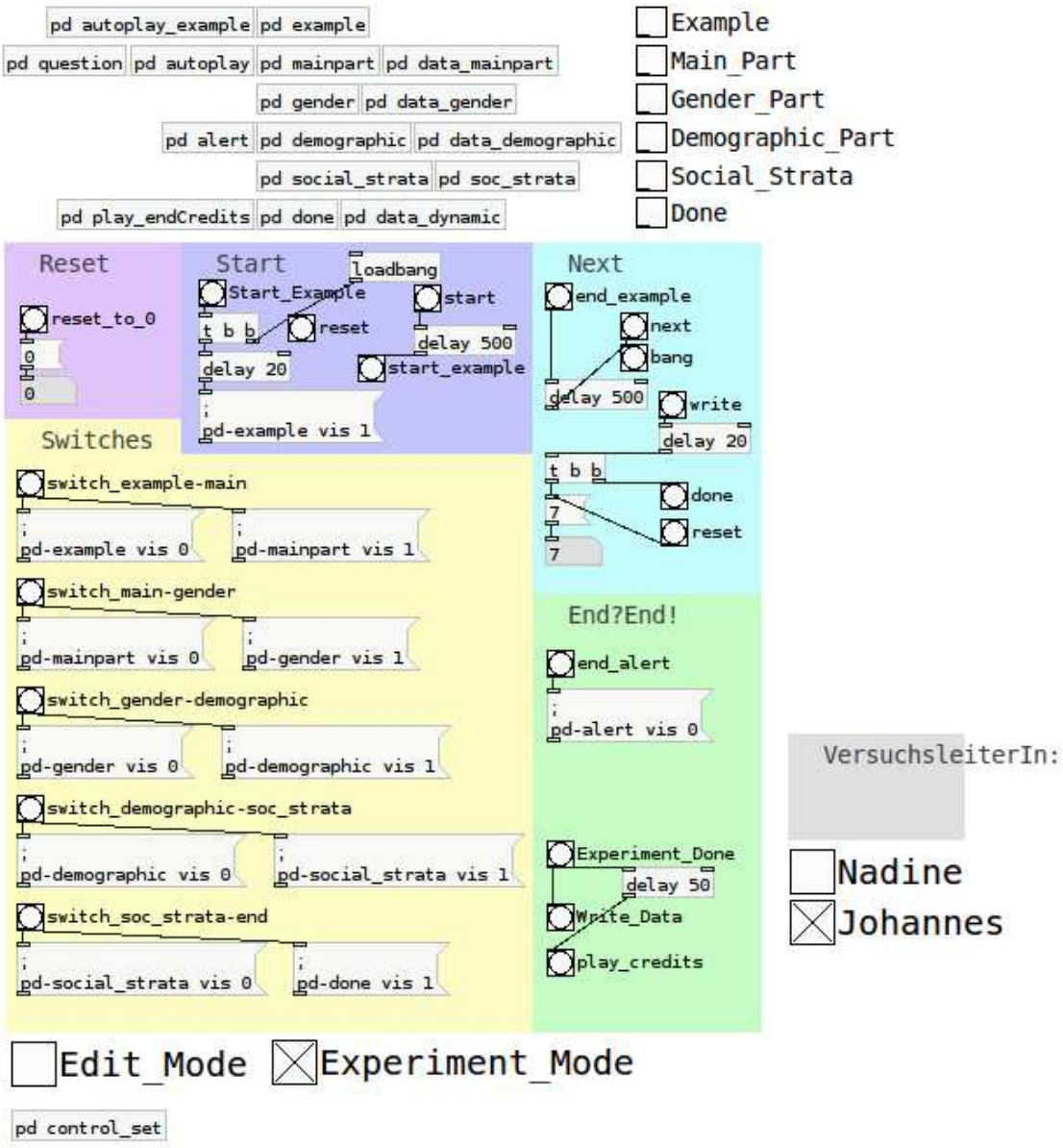


Figure C-1-6. Screenshot showing the control sub-patch.

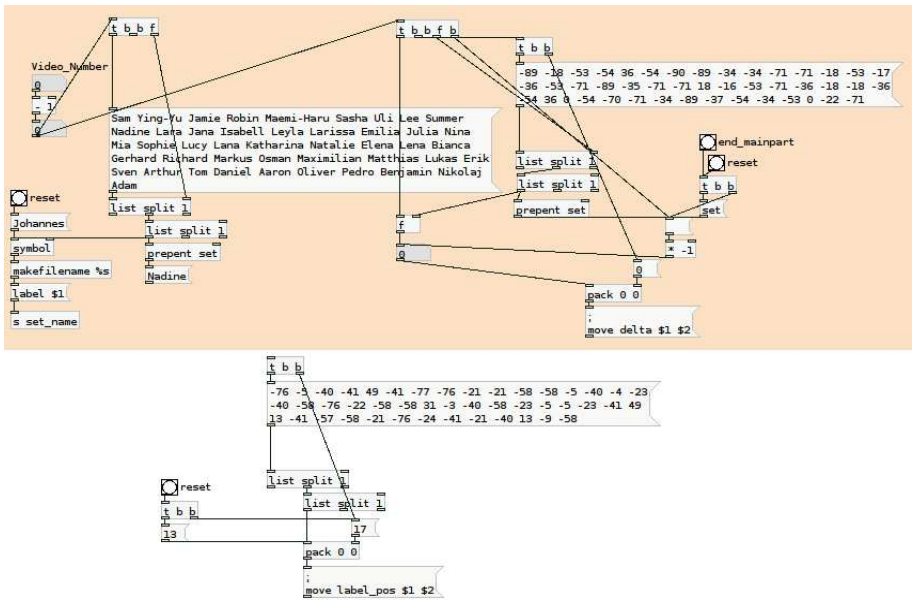


Figure C-1-7. Screenshot showing the control and background processes sub-patch of the main input screen.

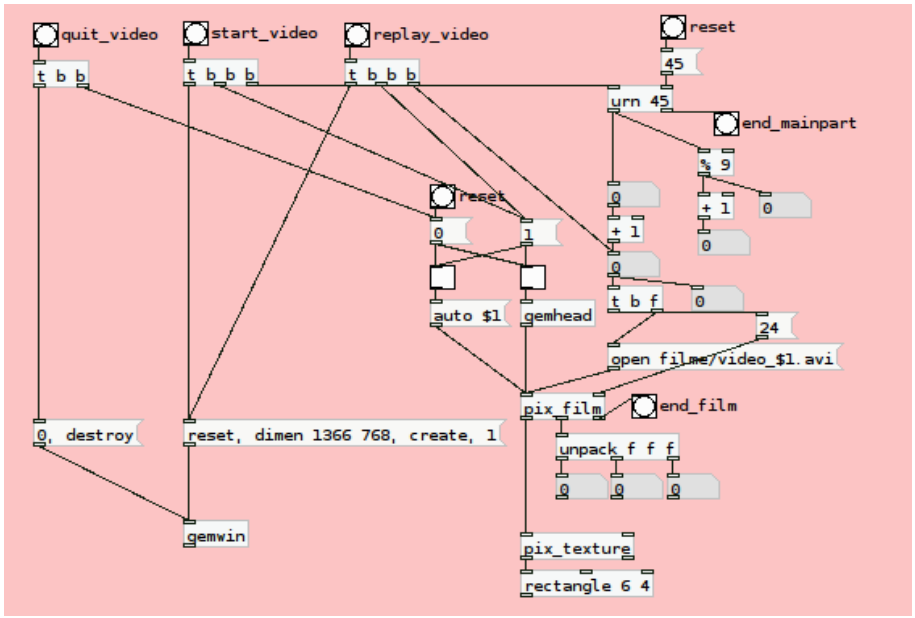


Figure C-1-8. Screenshot showing the background processes sub-patch of the video autoplay.

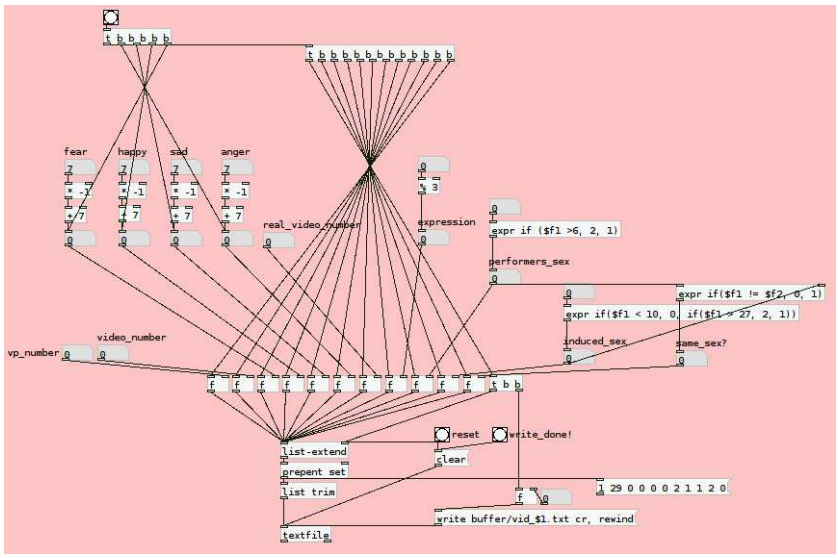


Figure C-1-9. Screenshot showing the background processes sub-patch of the temporary data save (writing the data into text files in the buffer file directory).

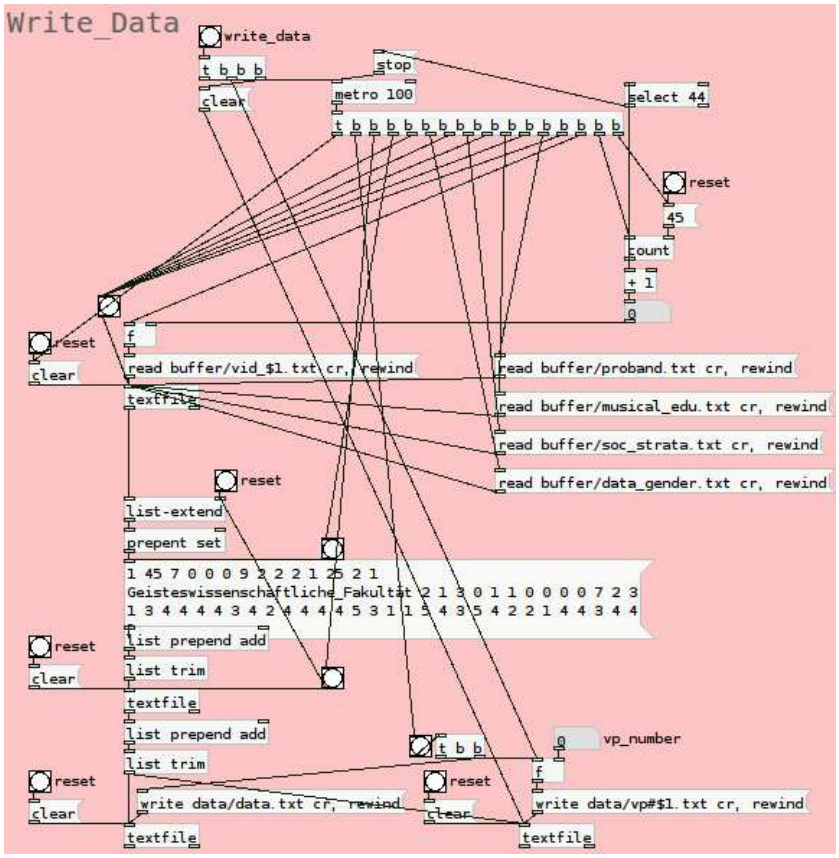


Figure C-1-10. Screenshot showing the background processes sub-patch of the full data set autosave (writing the data into text files in the data file directory).

Appendix C2 [Main Experiment – Videos]

Video Example

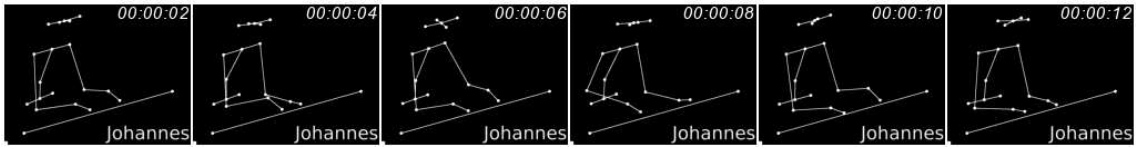


Figure C-2.1. Frames of the first example video of the main experiment (original video no.1, played by a female pianist with deadpan expression).

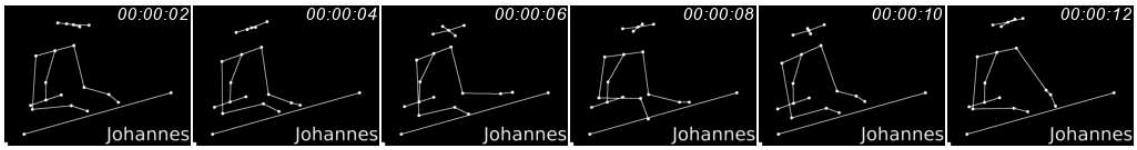


Figure C-2.2. Frames of the second example video of the main experiment (original video no.2, played by a female pianist with normal expression).

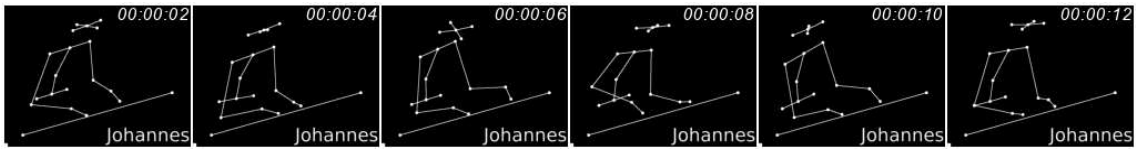


Figure C-2.3. Frames of the fourth example video of the main experiment (original video no.4, played by a female pianist with deadpan expression).

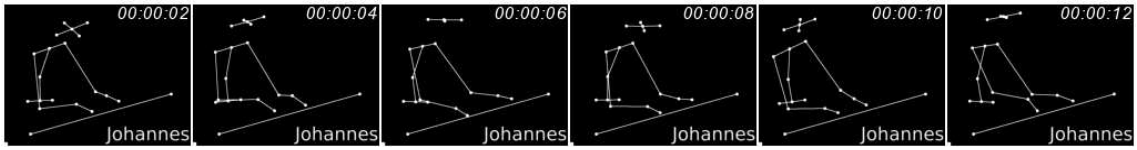


Figure C-2.4. Frames of the fifth example video of the main experiment (originally video no.5, played by a female pianist with normal expression).

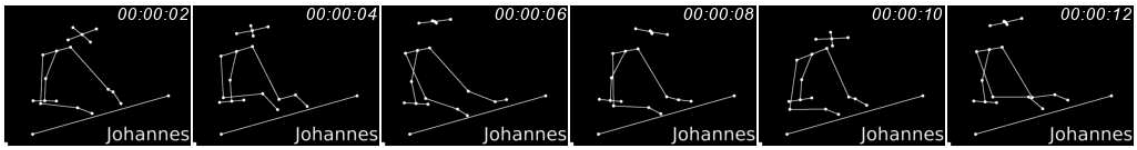


Figure C-2.5. Frames of the sixth example video of the main experiment (originally video no.6, played by a female pianist with exaggerated expression).

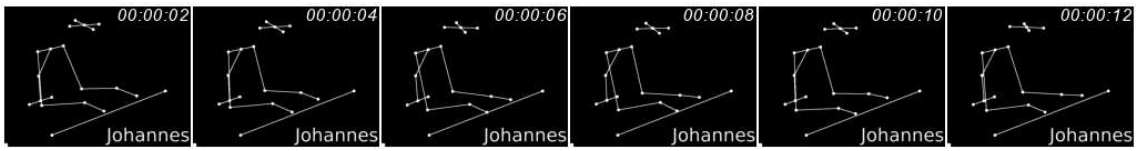


Figure C-2.6. Frames of the seventh example video of the main experiment (originally video no.7, played by a male pianist with deadpan expression).

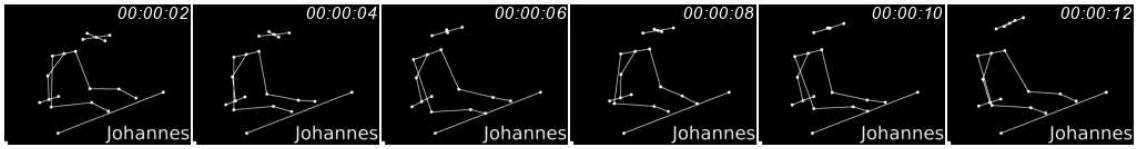


Figure C-2.7. Frames of the eight example video of the main experiment (originally video no.8, played by a male pianist with normal expression).

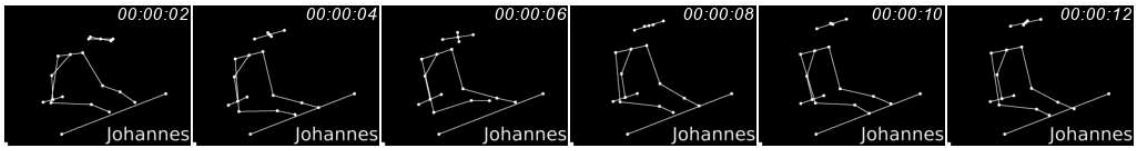


Figure C-2.8. Frames of the ninth example video of the main experiment (originally video no.9, played by a male pianist with exaggerated expression).

Videos with Neutral Gender Stimuli

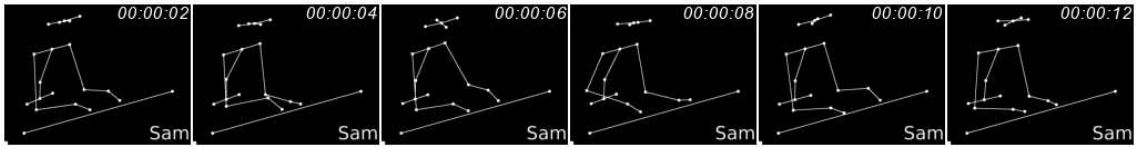


Figure C-2.9. Frames of the video clip labelled with 'Sam' (original video no.1, played by a female pianist with deadpan expression).

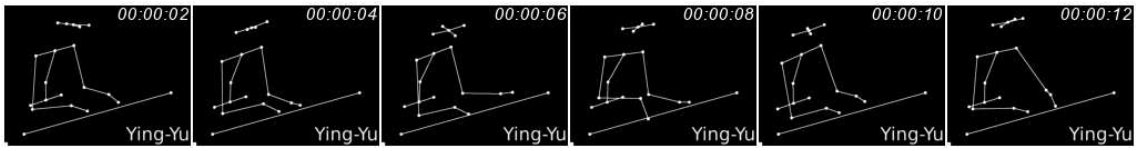


Figure C-2.10. Frames of the video clip labelled with 'Ying-Yu' (original video no.2, played by a female pianist with normal expression).

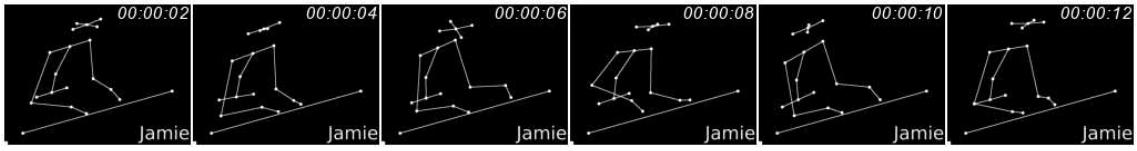


Figure C-2.11. Frames of the video clip labelled with 'Jamie' (original video no.3, played by a female pianist with exaggerated expression).

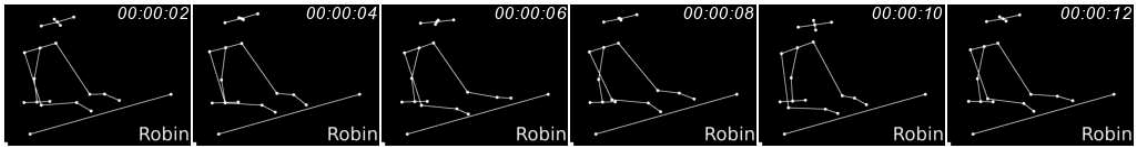


Figure C-2.12. Frames of the video clip labelled with 'Robin' (original video no.4, played by a female pianist with deadpan expression).

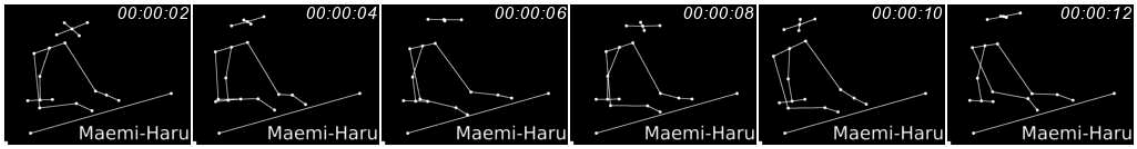


Figure C-2.13. Frames of the video clip labelled with 'Maemi-Haru' (originally video no.5, played by a female pianist with normal expression).

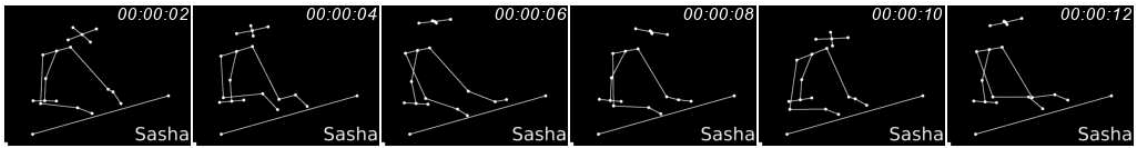


Figure C-2.14. Frames of the video clip labelled with 'Sasha' (originally video no.6, played by a female pianist with exaggerated expression).

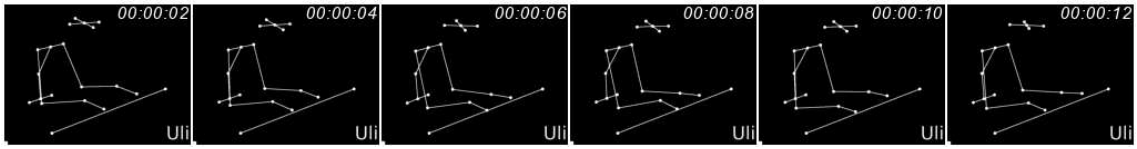


Figure C-2.15. Frames of the video clip labelled with 'Uli' (originally video no.7, played by a male pianist with deadpan expression).

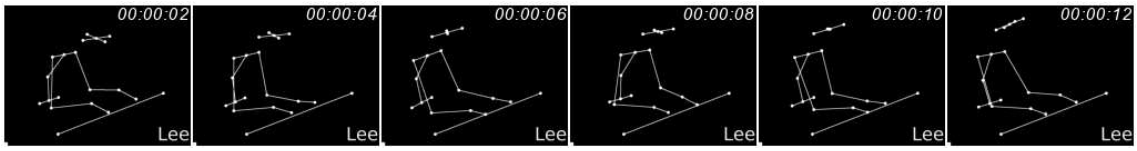


Figure C-2.16. Frames of the video clip labelled with 'Lee' (originally video no.8, played by a male pianist with normal expression).

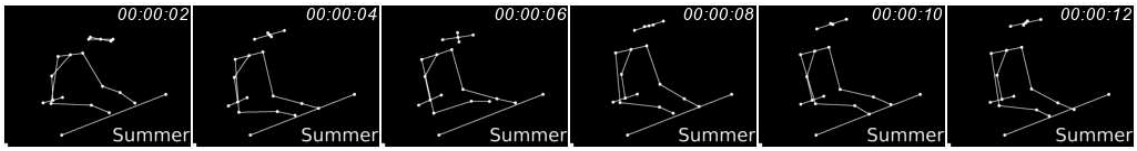


Figure C-2.17. Frames of the video clip labelled with 'Summer' (originally video no.9, played by a male pianist with exaggerated expression).

Videos with Female Gender Stimuli

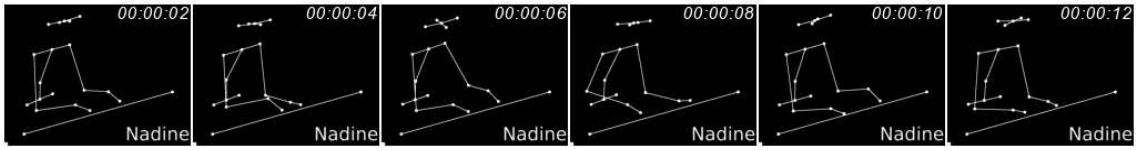


Figure C-2.18. Frames of the video clip labelled with 'Nadine' (original video no.1, played by a female pianist with deadpan expression).

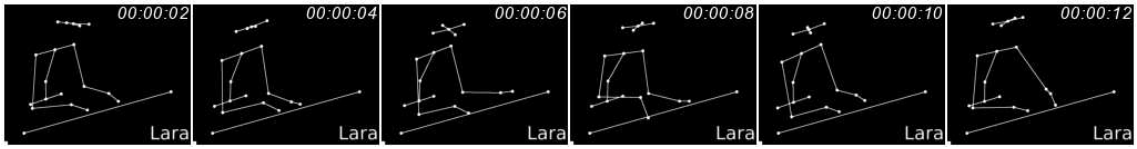


Figure C-2.19. Frames of the video clip labelled with 'Lara' (original video no.2, played by a female pianist with normal expression).

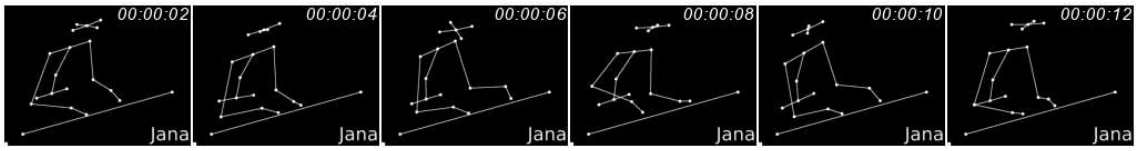


Figure C-2.20. Frames of the video clip labelled with 'Jana' (original video no.3, played by a female pianist with exaggerated expression).

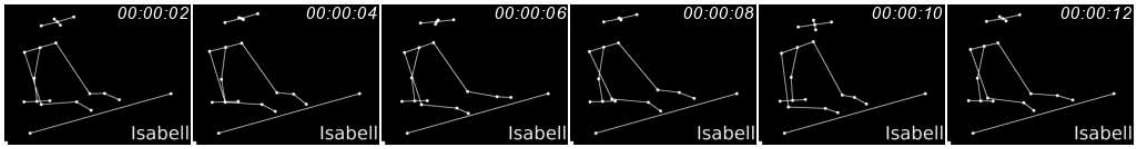


Figure C-2.21. Frames of the video clip labelled with 'Isabell' (original video no.4, played by a female pianist with deadpan expression).

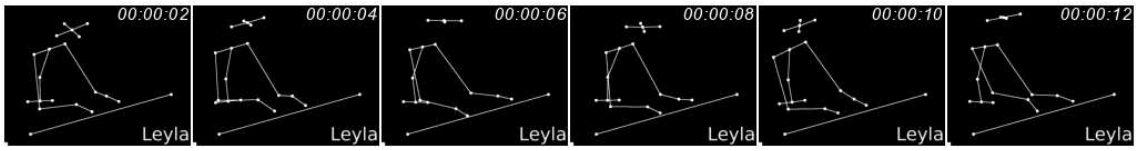


Figure C-2.22. Frames of the video clip labelled with 'Leyla' (originally video no.5, played by a female pianist with normal expression).

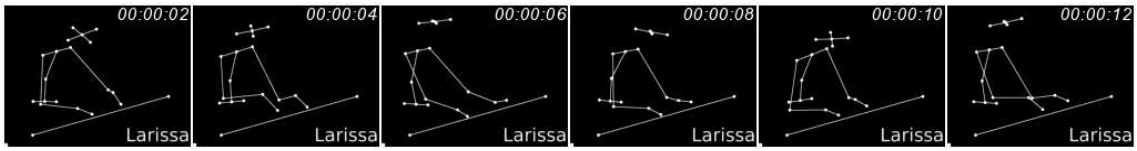


Figure C-2.23. Frames of the video clip labelled with 'Larissa' (originally video no.6, played by a female pianist with exaggerated expression).

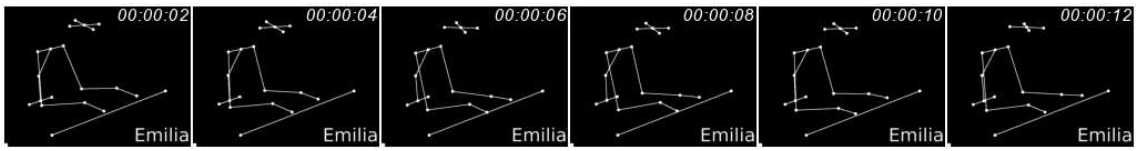


Figure C-2.24. Frames of the video clip labelled with 'Emilia' (originally video no.7, played by a male pianist with deadpan expression).

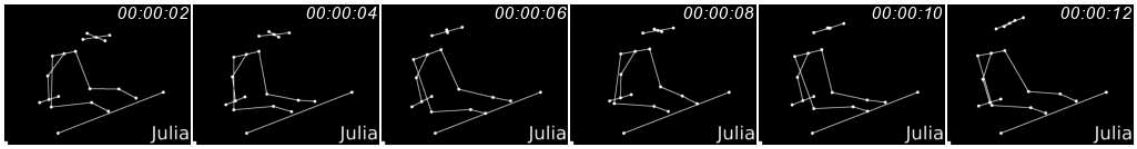


Figure C-2.25. Frames of the video clip labelled with 'Julia' (originally video no.8, played by a male pianist with normal expression).

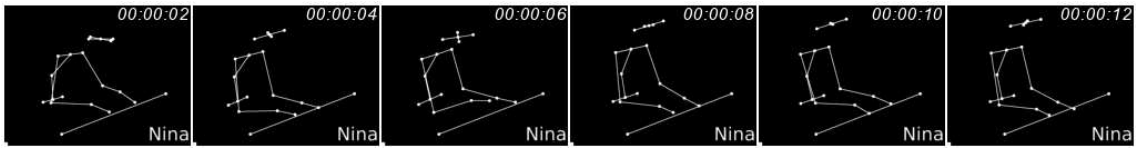


Figure C-2.26. Frames of the video clip labelled with 'Nina' (originally video no.9, played by a male pianist with exaggerated expression).

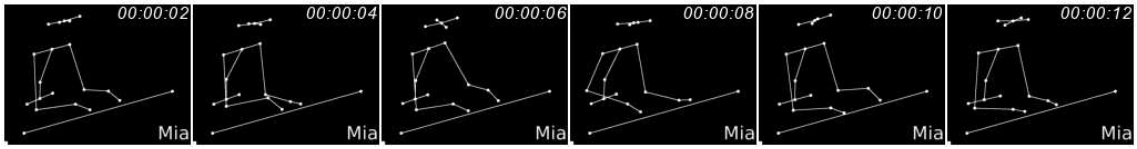


Figure C-2.27. Frames of the video clip labelled with 'Mia' (original video no.1, played by a female pianist with deadpan expression).

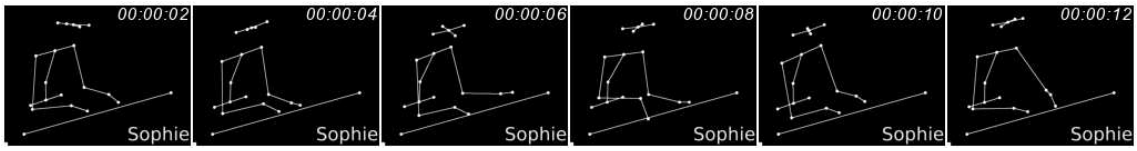


Figure C-2.28. Frames of the video clip labelled with 'Sophie' (original video no.2, played by a female pianist with normal expression).

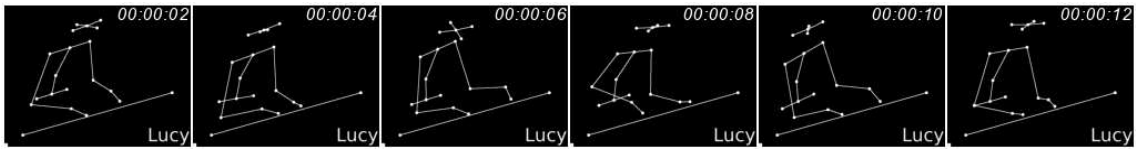


Figure C-2.29. Frames of the video clip labelled with 'Lucy' (original video no.3, played by a female pianist with exaggerated expression).

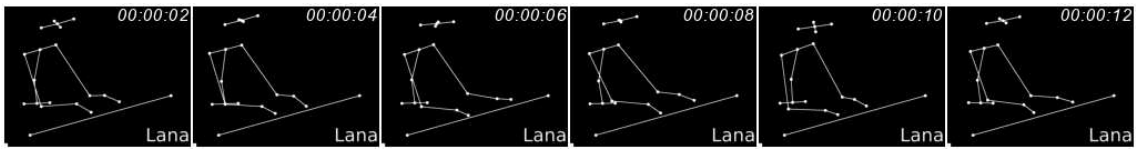


Figure C-2.30. Frames of the video clip labelled with 'Lana' (original video no.4, played by a female pianist with deadpan expression).

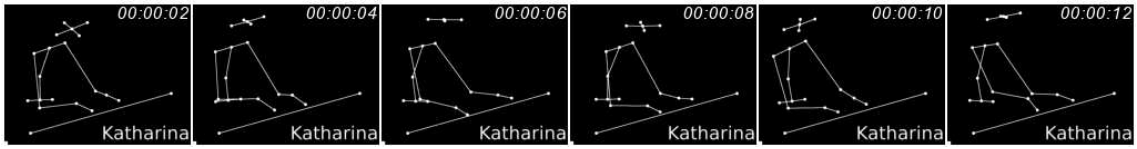


Figure C-2.31. Frames of the video clip labelled with 'Katharina' (originally video no.5, played by a female pianist with normal expression).

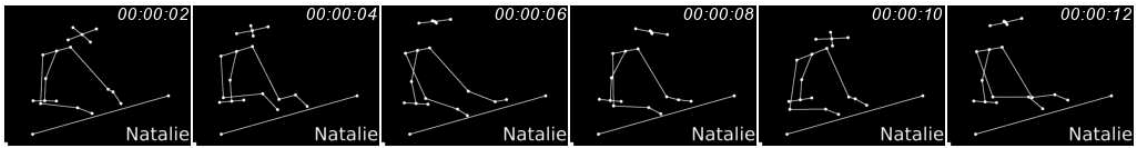


Figure C-2.32. Frames of the video clip labelled with 'Natalie' (originally video no.6, played by a female pianist with exaggerated expression).

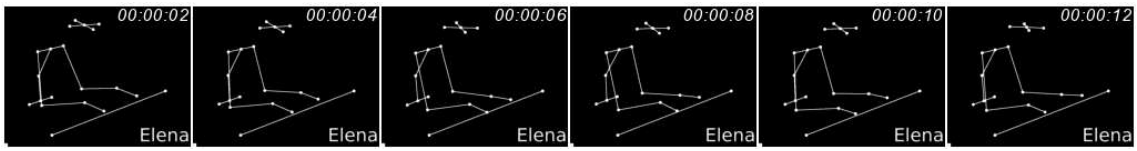


Figure C-2.33. Frames of the video clip labelled with 'Elena' (originally video no.7, played by a male pianist with deadpan expression).

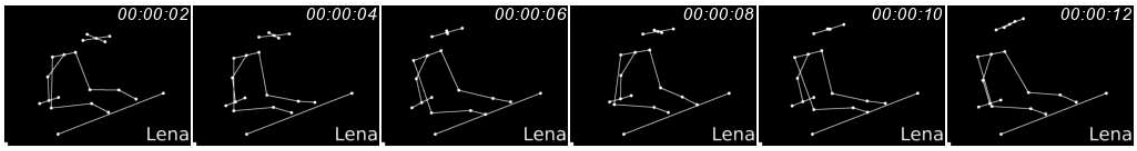


Figure C-2.34. Frames of the video clip labelled with 'Lena' (originally video no.8, played by a male pianist with normal expression).

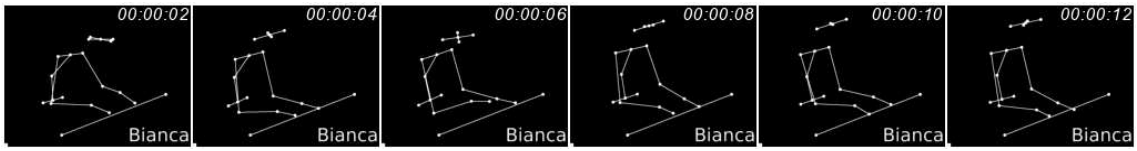


Figure C-2.35. Frames of the video clip labelled with 'Bianca' (originally video no.9, played by a male pianist with exaggerated expression).

Videos with Male Gender Stimuli

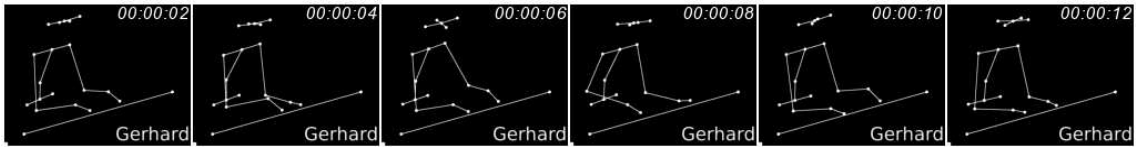


Figure C-2.36. Frames of the video clip labelled with 'Gerhard' (original video no.1, played by a female pianist with deadpan expression).

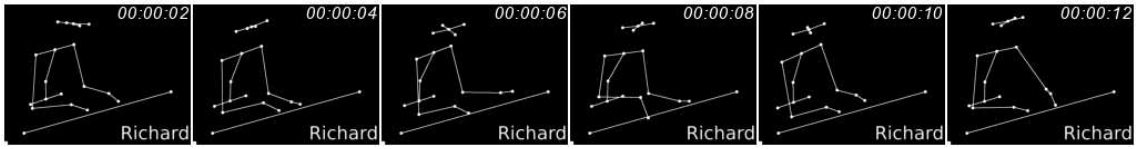


Figure C-2.37. Frames of the video clip labelled with 'Richard' (original video no.2, played by a female pianist with normal expression).

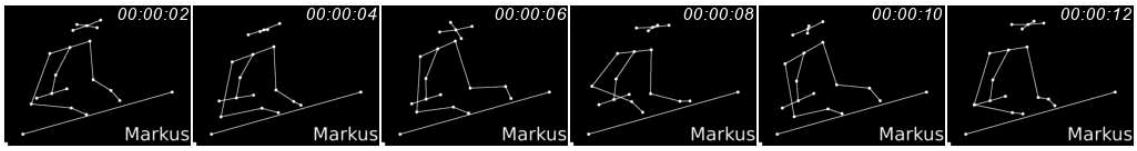


Figure C-2.38. Frames of the video clip labelled with 'Markus' (original video no.3, played by a female pianist with exaggerated expression).

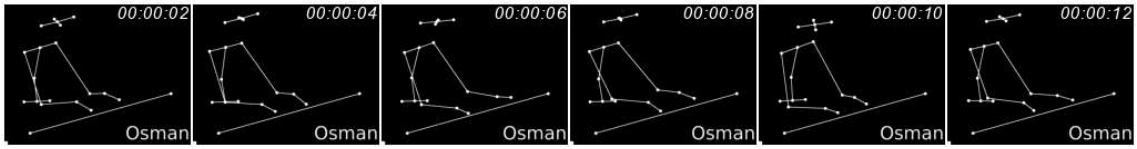


Figure C-2.39. Frames of the video clip labelled with 'Osman' (original video no.4, played by a female pianist with deadpan expression).

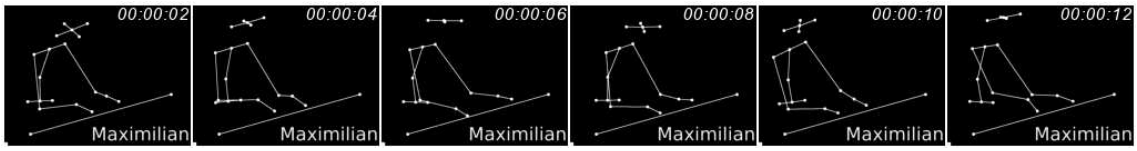


Figure C-2.40. Frames of the video clip labelled with 'Maximilian' (originally video no.5, played by a female pianist with normal expression).

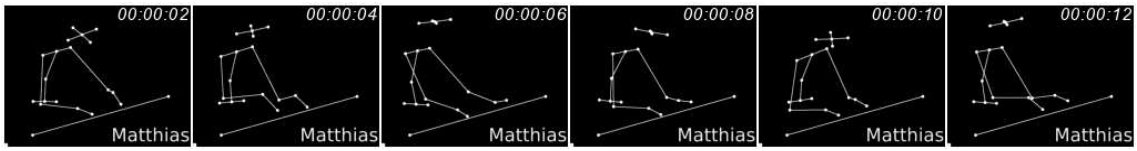


Figure C-2.41. Frames of the video clip labelled with 'Matthias' (originally video no.6, played by a female pianist with exaggerated expression).

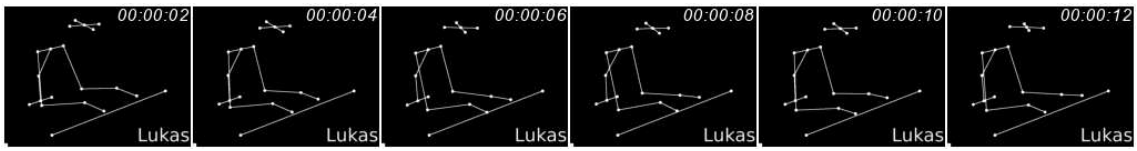


Figure C-2.42. Frames of the video clip labelled with 'Lukas' (originally video no.7, played by a male pianist with deadpan expression).

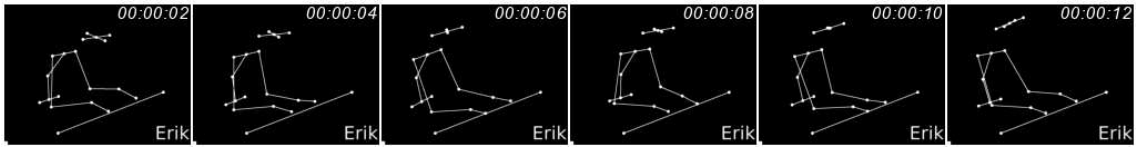


Figure C-2.43. Frames of the video clip labelled with 'Erik' (originally video no.8, played by a male pianist with normal expression).

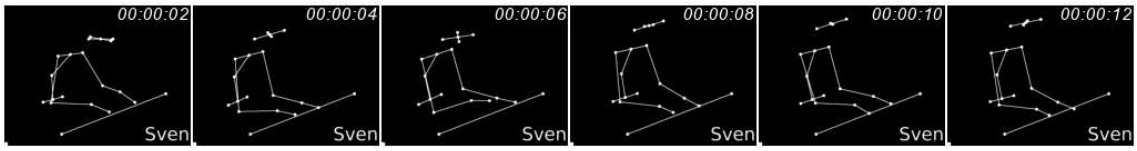


Figure C-2.44. Frames of the video clip labelled with 'Sven' (originally video no.9, played by a male pianist with exaggerated expression).

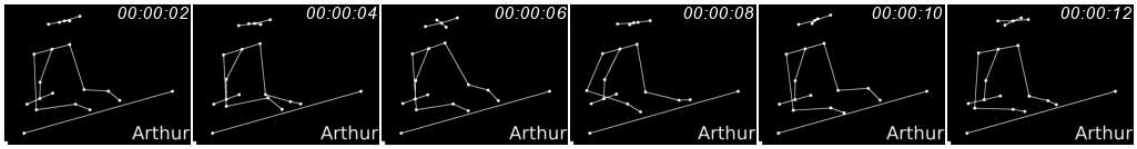


Figure C-2.45. Frames of the video clip labelled with 'Arthur' (original video no.1, played by a female pianist with deadpan expression).

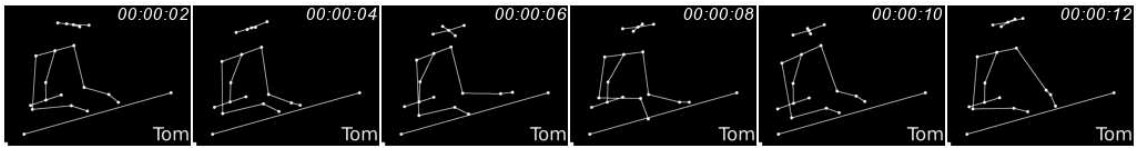


Figure C-2.46. Frames of the video clip labelled with 'Tom' (original video no.2, played by a female pianist with normal expression).

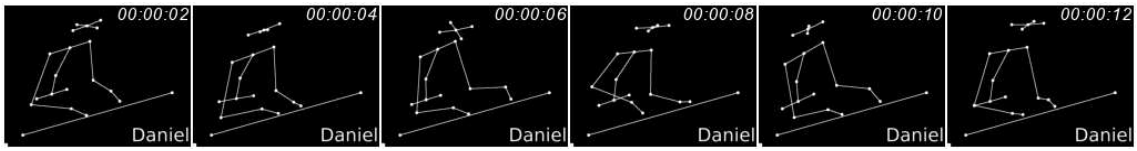


Figure C-2.47. Frames of the video clip labelled with 'Daniel' (original video no.3, played by a female pianist with exaggerated expression).

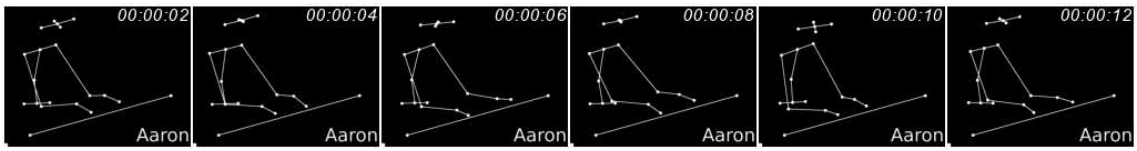


Figure C-2.48. Frames of the video clip labelled with 'Aaron' (original video no.4, played by a female pianist with deadpan expression).

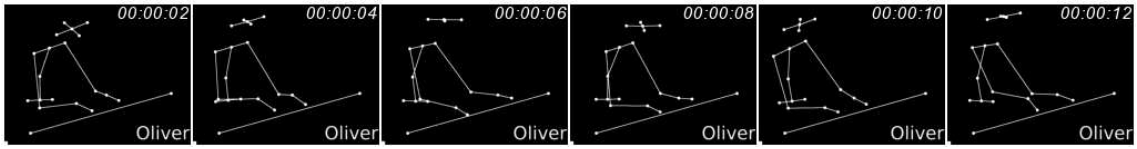


Figure C-2.49. Frames of the video clip labelled with 'Oliver' (originally video no.5, played by a female pianist with normal expression).

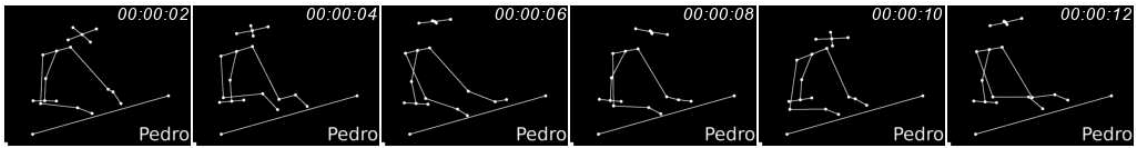


Figure C-2.50. Frames of the video clip labelled with 'Pedro' (originally video no.6, played by a female pianist with exaggerated expression).

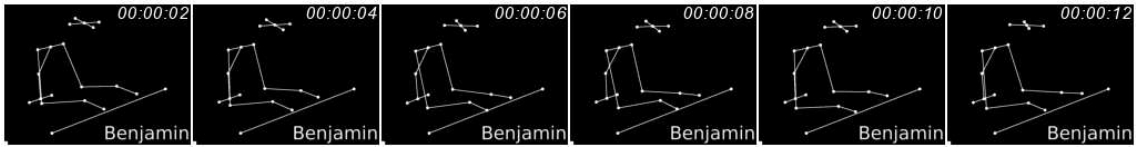


Figure C-2.51. Frames of the video clip labelled with 'Benjamin' (originally video no.7, played by a male pianist with deadpan expression).

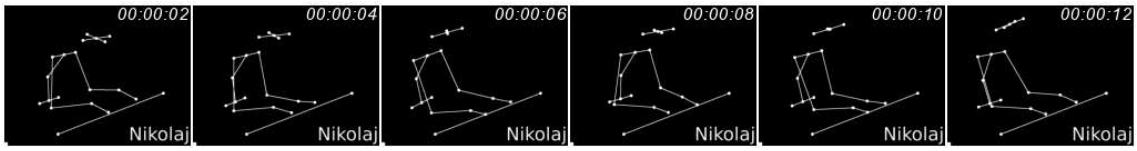


Figure C-2.52. Frames of the video clip labelled with 'Nikolaj' (originally video no.8, played by a male pianist with normal expression).

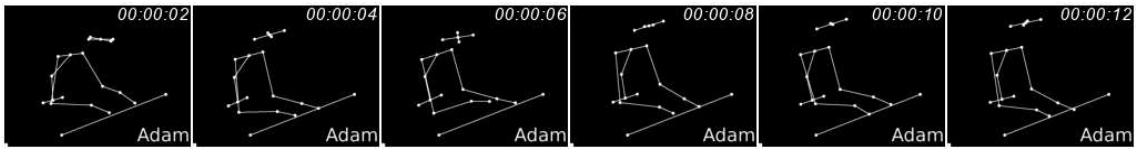


Figure C-2.53. Frames of the video clip labelled with 'Adam' (originally video no.9, played by a male pianist with exaggerated expression).

Appendix C3 [Main Experiment – Raw Data]

The raw data is available on the appendix CD only. All data is available in SPSS file format [*.sav], MS Excel 97/2003 format [*.xls], LibreOffice Calc (version 4.0) format [*.ods], or as plain text files [/content/experiments/main_study/data/..].

Table C-3.1

Demographic data including social gender, biological gender, age, highest education, duration of musical education, whether a key instrument is played or not, and the social strata.

	social gender[a]		biological		education			social
						musical		
VP#	E+/I+	E-/I-	gender[b]	age	highest[c]	years[d]	keys[e]	stratum[f]
1	3	1	2	31	3	3	1	-
2	1	1	1	20	1	0	-	4
3	2	2	2	27	1	3	0	-
4	4	3	1	27	1	3	1	-
5	3	3	2	36	1	3	0	-
6	2	4	1	18	1	3	1	5
7	4	4	1	23	1	3	0	4
8	4	4	1	24	2	3	0	5
9	2	4	2	26	2	0	-	-
10	4	2	2	21	1	3	0	4
11	3	3	1	22	1	3	1	2
12	3	4	1	21	1	2	0	5
13	3	2	1	21	1	3	1	5
14	1	1	1	25	1	1	0	3
15	2	3	2	25	2	3	1	4

(Table C-3.1 continues)

(Table C-3.1 continued)

	social gender[a]		biological		education			social
						musical		
VP#	E+/I+	E-/I-	gender[b]	age	highest[c]	years[d]	keys[e]	stratum[f]
16	1	3	1	47	3	3	1	-
17	3	2	1	21	1	3	0	2
18	2	1	2	27	3	3	1	-
19	3	3	1	24	2	1	1	4
20	2	2	2	20	1	3	1	5
21	1	3	1	30	3	3	1	-
22	2	2	2	30	3	3	1	-
23	4	3	1	22	1	3	0	5
24	3	2	2	27	2	3	0	-
25	4	4	1	18	2	3	1	4
26	4	1	1	21	1	3	0	5
27	2	1	1	24	2	2	1	5
28	4	3	2	22	1	0	-	4
29	1	1	1	22	1	3	1	2
30	3	4	2	20	1	3	0	4
31	3	2	1	24	1	3	1	5
32	1	3	1	47	3	3	1	-

Note. The social strata was not computed for participants older than 25 years; no values are available in the column for key instruments ('keys') for subjects who did not mention to be musical. All invalid (missing) values are marked with a minus (-).

a. 1 = feminine, 2 = masculine, 3 = androgynous, 4 = undifferentiated.

b. 1 = female, 2 = male.

c. 1 = matriculation, 2 = bachelor's degree, 3 = master's degree, 4 = doctorate

d. 0 = not musical, 1 = less than 1 year, 2 = 1 to 4 years, 3 = more than 4 years.

e. 0 = no, 1 = yes.

f. 2 = lower middle class, 3 = middle class, 4 = upper middle class, 5 = upper class.

Appendix C4 [Main Experiment – Tables]

The tables in appendix C contain all tables used in the paper as well as additional tables referred to in the paper in the main experiment section. All tables are digitally available in the MS Excel 97/2003 format [*.xls] and the LibreOffice Calc (version 4.0) format [*.ods] on the appendix CD [/content/files/appendix/03_main-experiment/tables/..].

Table C-4.1

Descriptions of the video clips used in the main experiment sorted by their video index number (names, index, assigned gender, original video number, pianist's real gender, and the intended expression).

	video index no.	assigned gender	original video no.	pianist's real gender	intended expression
Sam	1	neutral	1	female	deadpan
Ying-Yu	2	neutral	2	female	normal
Jamie	3	neutral	3	female	exaggerated
Robin	4	neutral	4	female	deadpan
Maemi-Haru	5	neutral	5	female	normal
Sasha	6	neutral	6	female	exaggerated
Uli	7	neutral	7	male	deadpan
Lee	8	neutral	8	male	normal
Summer	9	neutral	9	male	exaggerated
Nadine	10	female	1	female	deadpan
Lara	11	female	2	female	normal
Jana	12	female	3	female	exaggerated
Isabell	13	female	4	female	deadpan
Leyla	14	female	5	female	normal

(Table C-4.1 continues)

(Table C-4.1 continued)

	video index no.	assigned gender	original video no.	pianist's real gender	intended expression
Larissa	15	female	6	female	exaggerated
Emilia	16	female	7	male	deadpan
Julia	17	female	8	male	normal
Nina	18	female	9	male	exaggerated
Mia	19	female	1	female	deadpan
Sophie	20	female	2	female	normal
Lucy	21	female	3	female	exaggerated
Lana	22	female	4	female	deadpan
Katharina	23	female	5	female	normal
Natalie	24	female	6	female	exaggerated
Elena	25	female	7	male	deadpan
Lena	26	female	8	male	normal
Bianca	27	female	9	male	exaggerated
Gerhard	28	male	1	female	deadpan
Richard	29	male	2	female	normal
Markus	30	male	3	female	exaggerated
Osman	31	male	4	female	deadpan
Maximilian	32	male	5	female	normal
Matthias	33	male	6	female	exaggerated
Lukas	34	male	7	male	deadpan
Erik	35	male	8	male	normal
Sven	36	male	9	male	exaggerated
Arthur	37	male	1	female	deadpan
Tom	38	male	2	female	normal
Daniel	39	male	3	female	exaggerated
Aaron	40	male	4	female	deadpan

(Table C-4.1 continues)

(Table C-4.1 continued)

	video index no.	assigned gender	original video no.	pianist's real gender	intended expression
Oliver	41	male	5	female	normal
Pedro	42	male	6	female	exaggerated
Benjamin	43	male	7	male	deadpan
Nikolaj	44	male	8	male	normal
Adam	45	male	9	male	exaggerated

Descriptives

Table C-4.2

Descriptives (stated gender)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	neutral	288	1.28	1.812	.107	1.07	1.49
	female	576	1.33	1.829	.076	1.18	1.48
	male	576	1.26	1.727	.072	1.12	1.40
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	neutral	288	1.97	2.103	.124	1.72	2.21
	female	576	1.80	2.047	.085	1.63	1.97
	male	576	1.77	2.032	.085	1.61	1.94
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	neutral	288	1.74	2.015	.119	1.51	1.97
	female	576	1.88	2.110	.088	1.71	2.05
	male	576	1.92	2.085	.087	1.75	2.09
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	neutral	288	.58	1.304	.077	.43	.73
	female	576	.64	1.267	.053	.54	.74
	male	576	.61	1.258	.052	.51	.72
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.3

Descriptives (intended expression)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	deadpan	480	2.14	2.116	.097	1.95	2.33
	normal	480	1.19	1.655	.076	1.04	1.33
	exaggerated	480	.56	1.040	.047	.46	.65
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	deadpan	480	1.01	1.522	.069	.87	1.14
	normal	480	1.75	1.922	.088	1.58	1.92
	exaggerated	480	2.71	2.277	.104	2.51	2.91
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	deadpan	480	1.58	1.952	.089	1.41	1.76
	normal	480	2.13	2.170	.099	1.94	2.33
	exaggerated	480	1.89	2.083	.095	1.70	2.08
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	deadpan	480	.57	1.231	.056	.46	.68
	normal	480	.49	1.123	.051	.39	.59
	exaggerated	480	.79	1.423	.065	.66	.92
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.4

Descriptives (social gender categories according to positive personality traits)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	feminine	270	1.10	1.919	.117	.87	1.33
	masculine	360	1.46	1.914	.101	1.27	1.66
	androgynous	450	1.39	1.740	.082	1.23	1.55
	undifferentiated	360	1.14	1.570	.083	.98	1.31
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	feminine	270	1.74	2.080	.127	1.49	1.99
	masculine	360	1.70	2.053	.108	1.49	1.91
	androgynous	450	1.91	2.095	.099	1.71	2.10
	undifferentiated	360	1.90	1.976	.104	1.69	2.10
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	feminine	270	1.81	2.318	.141	1.53	2.09
	masculine	360	1.79	1.990	.105	1.58	2.00
	androgynous	450	2.08	1.996	.094	1.90	2.26
	undifferentiated	360	1.73	2.074	.109	1.52	1.95
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	feminine	270	.57	1.285	.078	.42	.73
	masculine	360	.60	1.218	.064	.47	.73
	androgynous	450	.80	1.437	.068	.66	.93
	undifferentiated	360	.44	1.046	.055	.33	.55
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.5

Descriptives (social gender categories according to negative personality traits)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	feminine	360	1.01	1.485	.078	.86	1.17
	masculine	315	1.30	1.693	.095	1.11	1.49
	androgynous	315	1.48	2.176	.123	1.23	1.72
	undifferentiated	450	1.38	1.741	.082	1.22	1.54
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	feminine	360	1.76	2.057	.108	1.54	1.97
	masculine	315	1.96	2.048	.115	1.73	2.18
	androgynous	315	1.63	2.086	.118	1.39	1.86
	undifferentiated	450	1.92	2.023	.095	1.73	2.11
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	feminine	360	1.53	1.894	.100	1.34	1.73
	masculine	315	1.92	1.889	.106	1.71	2.13
	androgynous	315	2.08	2.500	.141	1.80	2.36
	undifferentiated	450	1.95	2.003	.094	1.77	2.14
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	feminine	360	.54	1.217	.064	.42	.67
	masculine	315	.66	1.270	.072	.52	.80
	androgynous	315	.64	1.417	.080	.49	.80
	undifferentiated	450	.62	1.205	.057	.51	.74
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.6

Descriptives (weighted overall social gender categories)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	feminine	405	.86	1.371	.068	.73	.99
	masculine	405	1.54	1.840	.091	1.36	1.72
	androgynous	315	1.70	2.241	.126	1.45	1.95
	undifferentiated	315	1.12	1.509	.085	.95	1.29
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	feminine	405	1.78	2.066	.103	1.58	1.98
	masculine	405	1.82	1.983	.099	1.63	2.02
	androgynous	315	1.74	2.223	.125	1.50	1.99
	undifferentiated	315	1.96	1.945	.110	1.74	2.17
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	feminine	405	1.50	1.848	.092	1.32	1.68
	masculine	405	1.70	1.819	.090	1.53	1.88
	androgynous	315	2.58	2.510	.141	2.31	2.86
	undifferentiated	315	1.84	2.043	.115	1.61	2.06
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	feminine	405	.47	1.140	.057	.36	.58
	masculine	405	.76	1.275	.063	.63	.88
	androgynous	315	.85	1.586	.089	.67	1.03
	undifferentiated	315	.39	.979	.055	.29	.50
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.7

Descriptives (participants' biological gender)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	female	900	1.15	1.760	.059	1.04	1.27
	male	540	1.52	1.803	.078	1.37	1.67
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	female	900	1.72	2.062	.069	1.58	1.85
	male	540	2.00	2.026	.087	1.83	2.17
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	female	900	1.84	2.165	.072	1.70	1.98
	male	540	1.91	1.934	.083	1.75	2.08
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	female	900	.44	1.093	.036	.37	.51
	male	540	.91	1.476	.064	.78	1.03
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.8

Descriptives (musically educated – participants)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	yes	1305	1.28	1.796	.050	1.18	1.38
	no	135	1.43	1.669	.144	1.15	1.71
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	yes	1305	1.84	2.064	.057	1.73	1.95
	no	135	1.66	1.936	.167	1.33	1.99
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	yes	1305	1.92	2.102	.058	1.81	2.04
	no	135	1.36	1.794	.154	1.06	1.67
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	yes	1305	.60	1.271	.035	.53	.67
	no	135	.81	1.247	.107	.60	1.03
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.9

Descriptives (real gender of the pianist)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	female	960	1.21	1.675	.054	1.11	1.32
	male	480	1.45	1.977	.090	1.28	1.63
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	female	960	1.59	1.873	.060	1.48	1.71
	male	480	2.28	2.306	.105	2.07	2.48
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	female	960	2.20	2.149	.069	2.06	2.33
	male	480	1.21	1.768	.081	1.06	1.37
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	female	960	.39	.976	.031	.33	.46
	male	480	1.06	1.625	.074	.92	1.21
	Total	1440	.62	1.270	.033	.55	.68

Table C-4.10

Descriptives (social strata – participants)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	lower middle class	135	.87	1.463	.126	.63	1.12
	middle class	45	.29	.661	.099	.09	.49
	upper middle class	360	1.64	1.854	.098	1.45	1.84
	upper class	405	1.08	1.696	.084	.92	1.25
	Total	945	1.23	1.732	.056	1.12	1.34
happiness	lower middle class	135	1.40	1.821	.157	1.09	1.71
	middle class	45	.89	1.172	.175	.54	1.24
	upper middle class	360	1.89	2.131	.112	1.67	2.11
	upper class	405	1.79	2.158	.107	1.58	2.00
	Total	945	1.73	2.078	.068	1.60	1.86
sadness	lower middle class	135	1.88	1.985	.171	1.54	2.22
	middle class	45	.67	.853	.127	.41	.92
	upper middle class	360	1.91	2.172	.114	1.69	2.14
	upper class	405	2.01	2.172	.108	1.80	2.23
	Total	945	1.89	2.118	.069	1.76	2.03
anger	lower middle class	135	.73	1.411	.121	.49	.97
	middle class	45	.16	.424	.063	.03	.28
	upper middle class	360	.64	1.320	.070	.50	.78
	upper class	405	.37	.980	.049	.27	.46
	Total	945	.51	1.180	.038	.44	.59

Table C-4.11

Descriptives (highest level of education – participants)

						95% Confidence Interval for Mean	
		N	M	SD	SE	Lower Bound	Upper Bound
fear	matriculation	855	1.30	1.688	.058	1.18	1.41
	bachelor's degree	315	1.34	1.831	.103	1.14	1.54
	master's degree	270	1.22	2.017	.123	.98	1.46
	Total	1440	1.29	1.784	.047	1.20	1.38
happiness	matriculation	855	1.63	1.934	.066	1.50	1.76
	bachelor's degree	315	2.13	2.208	.124	1.89	2.38
	master's degree	270	2.06	2.167	.132	1.80	2.32
	Total	1440	1.82	2.053	.054	1.72	1.93
sadness	matriculation	855	1.78	1.991	.068	1.64	1.91
	bachelor's degree	315	1.97	2.158	.122	1.74	2.21
	master's degree	270	2.04	2.252	.137	1.77	2.31
	Total	1440	1.87	2.081	.055	1.76	1.98
anger	matriculation	855	.56	1.176	.040	.48	.64
	bachelor's degree	315	.61	1.322	.075	.46	.75
	master's degree	270	.81	1.468	.089	.64	.99
	Total	1440	.62	1.270	.033	.55	.68

ANOVA and Post-Hoc Tests

Table C-4.12

Analysis of Variance (stated gender)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	1.626	2	.813	.255	.775
	Within Groups	4580.290	1437	3.187		
	Total	4581.916	1439			
happiness	Between Groups	7.590	2	3.795	.901	.407
	Within Groups	6054.899	1437	4.214		
	Total	6062.489	1439			
sadness	Between Groups	6.426	2	3.213	.742	.477
	Within Groups	6225.767	1437	4.332		
	Total	6232.194	1439			
anger	Between Groups	.800	2	.400	.248	.781
	Within Groups	2321.366	1437	1.615		
	Total	2322.166	1439			

Table C-4.13

Analysis of Variance (intended expression)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	606.739	2	303.369	109.666	.000
	Within Groups	3975.177	1437	2.766		
	Total	4581.916	1439			
happiness	Between Groups	700.760	2	350.380	93.906	.000
	Within Groups	5361.729	1437	3.731		
	Total	6062.489	1439			
sadness	Between Groups	73.529	2	36.765	8.578	.000
	Within Groups	6158.665	1437	4.286		
	Total	6232.194	1439			
anger	Between Groups	22.268	2	11.134	6.957	.001
	Within Groups	2299.898	1437	1.600		
	Total	2322.166	1439			

Table C-4.13a

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'fear'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			2.14	1.95	2.33
normal	.000*	1.000		1.19	1.04	1.33
exaggerated	.000*	.000*	1.000	.56	.46	.65
Total				1.29	1.20	1.38

*p < .001

Table C-4.13b

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'happiness'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			1.01	.87	1.14
normal	.000*	1.000		1.75	1.58	1.92
exaggerated	.000*	.000*	1.000	2.71	2.51	2.91
Total				1.82	1.72	1.93

*p < .001

Table C-4.13c

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'sadness'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			1.58	1.41	1.76
normal	.000*	1.000		2.13	1.94	2.33
exaggerated	.053	.167	1.000	1.89	1.70	2.08
Total				1.87	1.76	1.98

*p < .001

Table C-4.13d

Tukey post-hoc comparisons of the expression intended to be played by the pianist for the dependent variable 'anger'.

					95% Confidence Interval for Mean	
	deadpan	normal	exaggerated	M	Lower Bound	Upper Bound
deadpan	1.000			.57	.46	.68
normal	.613	1.000		.49	.39	.59
exaggerated	.022*	.001**	1.000	.79	.66	.92
Total				.62	.55	.68

*p < .05, **p < .01

Table C-4.14

Analysis of Variance (social gender categories according to positive personality traits)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	32.652	3	10.884	3.436	.016
	Within Groups	4549.264	1436	3.168		
	Total	4581.916	1439			
happiness	Between Groups	12.575	3	4.192	.995	.394
	Within Groups	6049.914	1436	4.213		
	Total	6062.489	1439			
sadness	Between Groups	30.269	3	10.090	2.336	.072
	Within Groups	6201.925	1436	4.319		
	Total	6232.194	1439			
anger	Between Groups	26.375	3	8.792	5.499	.001
	Within Groups	2295.791	1436	1.599		
	Total	2322.166	1439			

Table C-4.15

Analysis of Variance (social gender categories according to negative personality traits)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	42.782	3	14.261	4.512	.004
	Within Groups	4539.134	1436	3.161		
	Total	4581.916	1439			
happiness	Between Groups	23.705	3	7.902	1.879	.131
	Within Groups	6038.783	1436	4.205		
	Total	6062.489	1439			
sadness	Between Groups	58.542	3	19.514	4.539	.004
	Within Groups	6173.652	1436	4.299		
	Total	6232.194	1439			
anger	Between Groups	2.838	3	.946	.586	.624
	Within Groups	2319.328	1436	1.615		
	Total	2322.166	1439			

Table C-4.16

Analysis of Variance (weighted overall social gender categories)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	162.679	3	54.226	17.620	.000
	Within Groups	4419.237	1436	3.077		
	Total	4581.916	1439			
happiness	Between Groups	8.654	3	2.885	.684	.562
	Within Groups	6053.835	1436	4.216		
	Total	6062.489	1439			
sadness	Between Groups	227.236	3	75.745	18.113	.000
	Within Groups	6004.957	1436	4.182		
	Total	6232.194	1439			
anger	Between Groups	49.841	3	16.614	10.499	.000
	Within Groups	2272.325	1436	1.582		
	Total	2322.166	1439			

Table C-4.16a

Tukey post-hoc comparisons of the weighted overall social gender categories for the dependent variable 'fear'.

						95% Confidence Interval for Mean	
	fem.	masc.	andr.	undif.	M	Lower Bound	Upper Bound
feminine	1.000				.86	.73	.99
masculine	.000**	1.000			1.54	1.36	1.72
androgynous	.000**	.641	1.000		1.70	1.45	1.95
undifferentiated	.195	.007*	.000**	1.000	1.12	.95	1.29
Total					1.29	1.20	1.38

*p < .01, **p < .001

Table C-4.16b

Tukey post-hoc comparisons of the weighted overall social gender categories for the dependent variable 'sadness'.

						95% Confidence Interval for Mean	
	fem.	masc.	andr.	undif.	M	Lower Bound	Upper Bound
feminine	1.000				1.50	1.32	1.68
masculine	.494	1.000			1.70	1.53	1.88
androgynous	.000*	.000*	1.000		2.58	2.31	2.86
undifferentiated	.126	.818	.000*	1.000	1.84	1.61	2.06
Total					1.87	1.76	1.98

*p < .001

Table C-4.16c

Tukey post-hoc comparisons of the weighted overall social gender categories for the dependent variable 'anger'.

						95% Confidence Interval for Mean	
	fem.	masc.	andr.	undif.	M	Lower Bound	Upper Bound
feminine	1.000				.47	.36	.58
masculine	.006*	1.000			.76	.63	.88
androgynous	.000**	.760	1.000		.85	.67	.103
undifferentiated	.855	.001*	.000**	1.000	.39	.29	.50
Total					.62	.55	.68

*p < .01, **p < .001

Table C-4.17

Analysis of Variance (participants' biological gender)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	45.650	1	45.650	14.471	.000
	Within Groups	4536.266	1438	3.155		
	Total	4581.916	1439			
happiness	Between Groups	26.741	1	26.741	6.371	.012
	Within Groups	6035.748	1438	4.197		
	Total	6062.489	1439			
sadness	Between Groups	1.689	1	1.689	.390	.533
	Within Groups	6230.505	1438	4.333		
	Total	6232.194	1439			
anger	Between Groups	73.617	1	73.617	47.080	.000
	Within Groups	2248.549	1438	1.564		
	Total	2322.166	1439			

Table C-4.18

Analysis of Variance (musically educated – participants)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	2.807	1	2.807	.881	.348
	Within Groups	4579.109	1438	3.184		
	Total	4581.916	1439			
happiness	Between Groups	3.956	1	3.956	.939	.333
	Within Groups	6058.533	1438	4.213		
	Total	6062.489	1439			
sadness	Between Groups	38.108	1	38.108	8.847	.003
	Within Groups	6194.085	1438	4.307		
	Total	6232.194	1439			
anger	Between Groups	5.808	1	5.808	3.606	.058
	Within Groups	2316.358	1438	1.611		
	Total	2322.166	1439			

Table C-4.19

Analysis of Variance (real gender of the pianist)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	18.850	1	18.850	5.940	.015
	Within Groups	4563.066	1438	3.173		
	Total	4581.916	1439			
happiness	Between Groups	148.967	1	148.967	36.225	.000
	Within Groups	5913.522	1438	4.112		
	Total	6062.489	1439			
sadness	Between Groups	308.113	1	308.113	74.791	.000
	Within Groups	5924.081	1438	4.120		
	Total	6232.194	1439			
anger	Between Groups	144.006	1	144.006	95.071	.000
	Within Groups	2178.160	1438	1.515		
	Total	2322.166	1439			

Table C-4.20

Analysis of Variance (highest level of education – participants)

		Sum of Squares	df	Mean Square	F	Sig.
fear	Between Groups	2.053	2	1.026	.322	.725
	Within Groups	4579.863	1437	3.187		
	Total	4581.916	1439			
happiness	Between Groups	76.354	2	38.177	9.165	.000
	Within Groups	5986.135	1437	4.166		
	Total	6062.489	1439			
sadness	Between Groups	18.961	2	9.480	2.193	.112
	Within Groups	6213.233	1437	4.324		
	Total	6232.194	1439			
anger	Between Groups	12.964	2	6.482	4.034	.018
	Within Groups	2309.202	1437	1.607		
	Total	2322.166	1439			

Linear Regression

Table C-4.21

R-Squared Table for the Model 'intended expression, stated gender'.

	R[a]	R Square	Adjusted R Square	SE of the Estimate
fear	.362	.131	.130	1.665
happiness	.341	.116	.115	1.931
sadness	.068	.005	.003	2.078
anger	.070	.005	.004	1.268

a. Predictors: (Constant), intended expression, stated gender.

Table C-4.22

Linear regression analyses with intended expression and stated gender as independent variables for fear, happiness, sadness, and anger as dependent variables.

		Unstandardised Coefficients		Standardised Coefficients		
		B	SE	Beta	t	Sig.
fear	(Constant)	2.109	.099		21.346	.000***
	intended expression	-.790	.054	-.361	-14.695	.000***
	stated gender	-.023	.059	-.009	-.385	.700
happiness	(Constant)	1.074	.115		9.371	.000***
	intended expression	.852	.062	.339	13.671	.000***
	stated gender	-.087	.068	-.032	-1.273	.203
sadness	(Constant)	1.614	.123		13.091	.000***
	intended expression	.155	.067	.061	2.315	.021*
	stated gender	.083	.073	.030	1.132	.258
anger	(Constant)	.494	.075		6.562	.000***
	intended expression	.108	.041	.070	2.647	.008**
	stated gender	.013	.045	.007	.283	.777

*p < .05, **p < .01, ***p < .001

Appendix C5 [Main Experiment – Figures]

The figures in appendix C contain all figures used in the paper as well as additional figures referred to in the paper in the main experiment section. All figures are digitally available in the Portable Network Graphic format [*.png] on the appendix CD [/content/files/appendix/03_main-experiment/figures/..].

Gender distribution

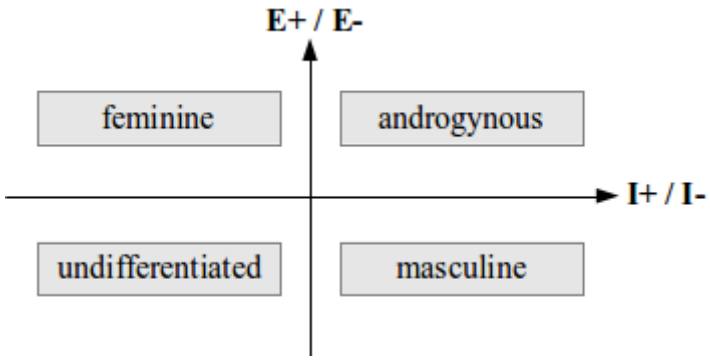


Figure 23. Interpretation of the 2-dimensional gender systems (combined graphic for positive and negative scaling).

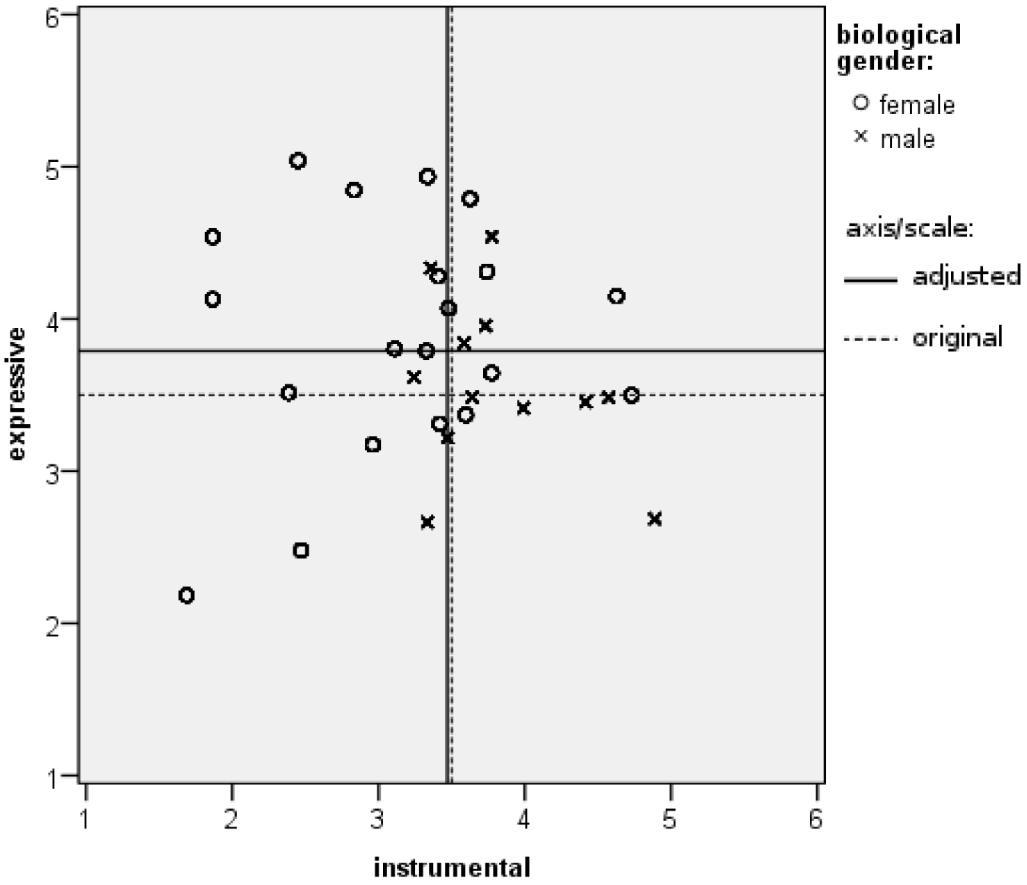


Figure 24. Gender distribution of the subjects depending on the weighted overall expressive and instrumental scale (weighted means).

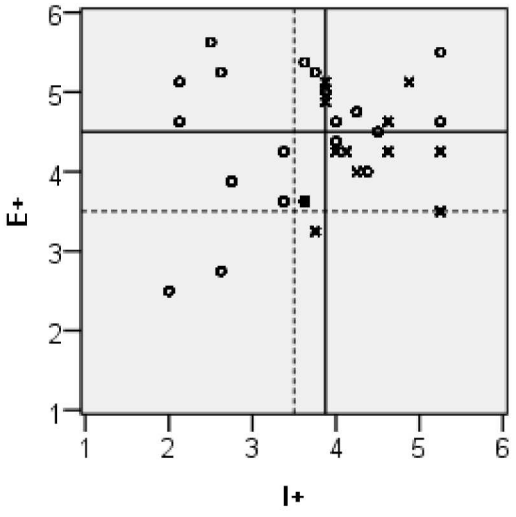


Figure 15a. Gender distribution depending on a positive expressive and instrumental scale.

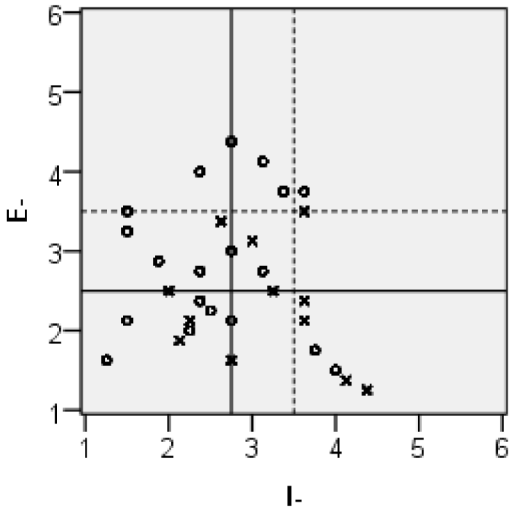


Figure 15b. Gender distribution depending on a negative expressive and instrumental scale.

Comparison of mean ratings

Figure 25. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the playing conditions of the stated gender 'neutral', 'female', and 'male' (error bars: 95% confidence interval).

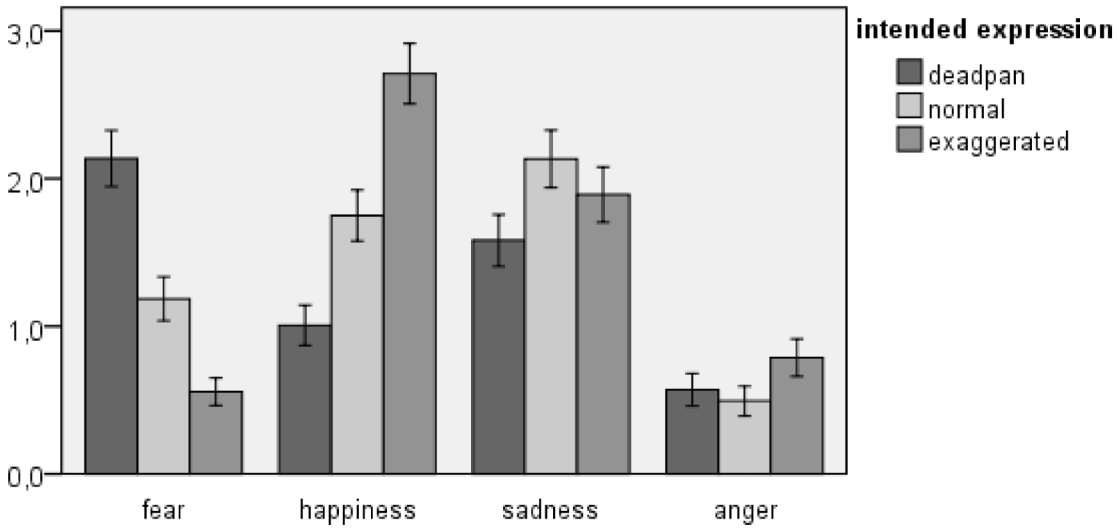


Figure 26. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the playing conditions 'deadpan', 'normal', and 'exaggerated' (error bars: 95% confidence interval).

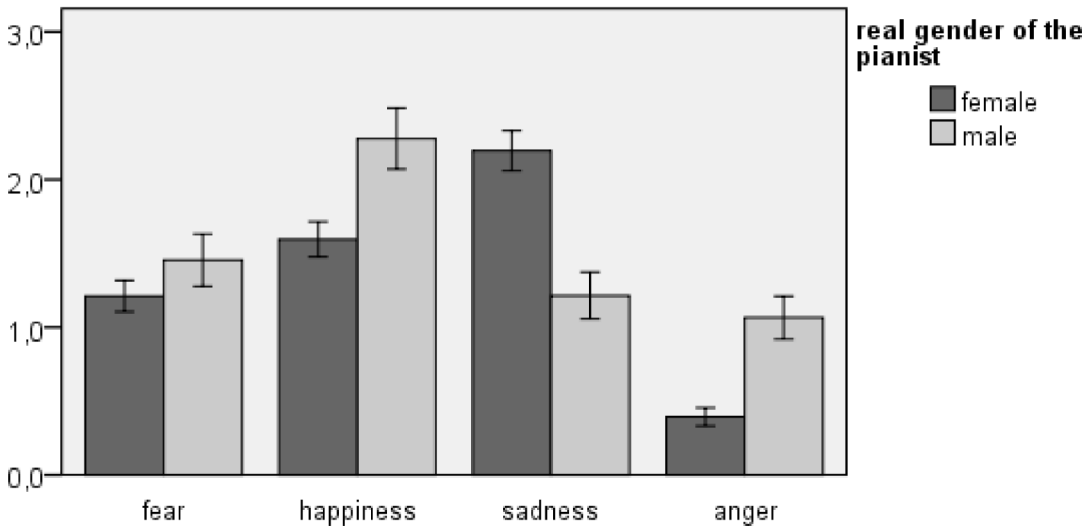


Figure 27. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the real biological gender of the pianist (error bars: 95% confidence interval).

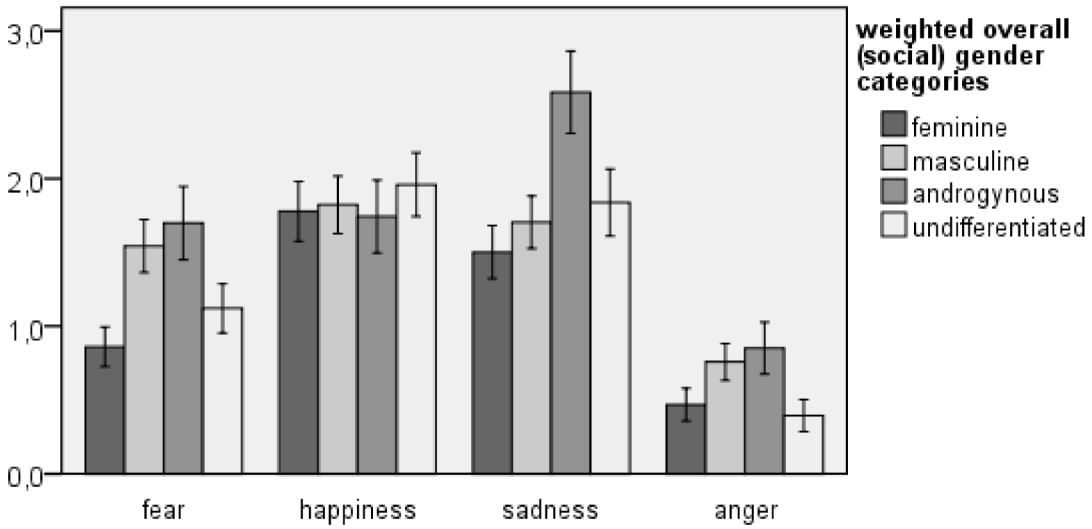


Figure 28. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the gender categories 'feminine', 'masculine', 'androgynous' and 'undifferentiated' (error bars: 95% confidence interval).

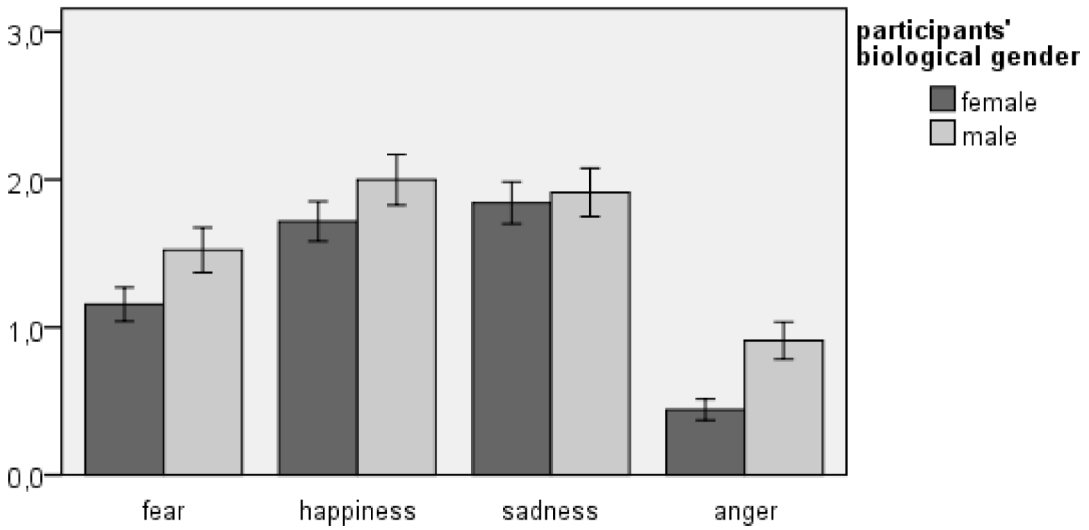


Figure 29. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the participants' biological gender (error bars: 95% confidence interval).

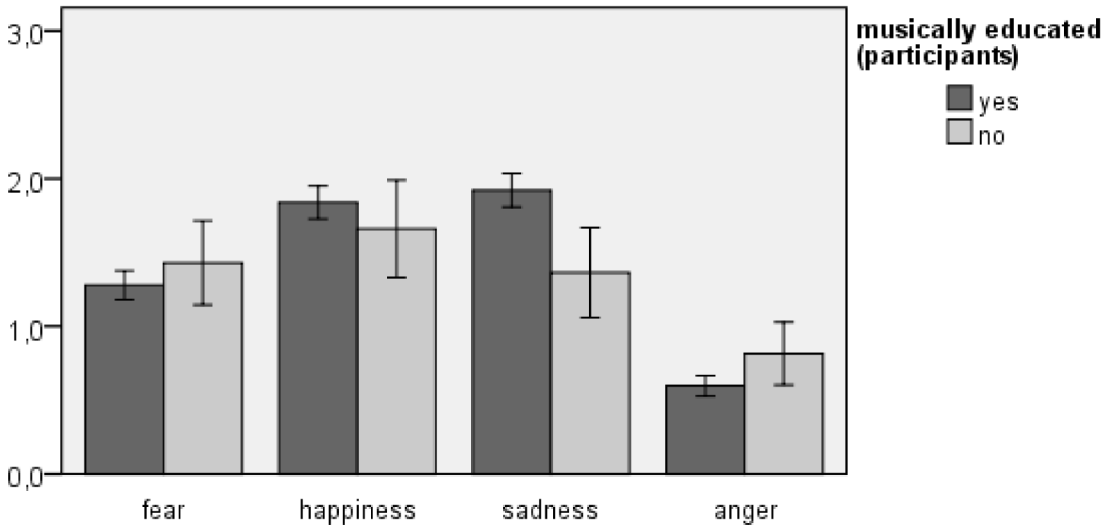


Figure 30. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the participants' musicality (error bars: 95% confidence interval).

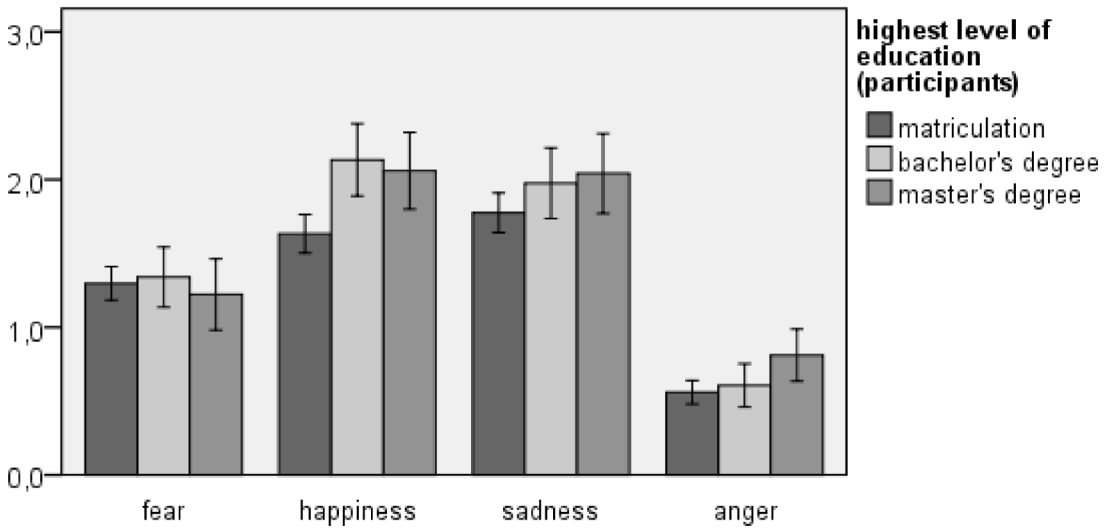


Figure 31. Comparison of mean ratings of all four emotions (fear, happiness, sadness, and anger) among the participants' highest level of education (error bars: 95% confidence interval).

Appendix D [Appendix CD – Table of Contents]

A full Table of Contents is also available on the appendix CD in Adobe Portable Document Format [index.pdf], a plain text file [index.txt], as well as a HTML5/CSS3-based Table of Contents [index.html]. The HTML-based index is also available online at http://www.johannes-lehner.at/masterthesis/ including the digital appendix.

(the folder-structure is written bold, important files are included in italic font-style)

appendix-CD-drive:

- *index.html* [Table of Contents]
- *index.pdf* [Table of Contents]
- *index.txt* [Table of Contents]
- **\content**
  - *masterthesis_lehner-2013.pdf* [Full Version with additional appendices]
  - **\experiments**
    - **\main_study**
      - *main_experiment.pd* [pd-patch used to conduct the experiment]
      - **\buffer** [includes example files only, used as temporary storage for the pd-patch]
      - **\data** [includes all raw data of the main experiment]
      - **\filme** [includes all videos of the main experiment]
    - **\pre-experiment_#1**
      - *feedback_questionnaire.pdf* [Demographic Data/Feedback Questionnaire]
      - *instructions.pdf* [instructions to the first pre-experiment]
      - *pre-experiment_#1.pd* [pd-patch used to conduct the first pre-experiment]
      - **\data** [includes all raw data of the first pre-experiment]
      - **\filme** [includes all videos of the first pre-experiment]
    - **\pre-experiment_#2**
      - *demographic.pdf* [Demographic Data Sheet]
      - *instructions.pdf* [instructions to the second pre-experiment]
      - **\data** [includes all raw data of the second pre-experiment]
      - **\questionnaires** [includes all empty name lists used in the first pre-experiment]

- **\files**
  - **\appendix**
    - **\01_pre-experiment1**
      - **\figures** [includes all figures, as well as screenshots of the pd-patch, of the first pre-experiment]
      - **\tables** [includes all tables of the first pre-experiment]
    - **\02_pre-experiment2**
      - **\figures** [includes all figures, as well as screenshots of the pd-patch, of the second pre-experiment]
      - **\tables** [includes all tables of the second pre-experiment]
    - **\03_main-experiment**
      - **\figures** [includes all figures, as well as screenshots of the pd-patch, of the main experiment]
      - **\tables** [includes all tables of the main experiment]
  - **\pd-extended_0.43.1** [installation files of the used pd-version]
- **\html** [includes files for 'index.html' only].