



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

A multi-method approach to assess socially desirable response bias in
personality questionnaires

verfasst von / submitted by

Philipp Gewessler, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master Of Science (MSc)

Wien, 2016 / Vienna, 2016

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 840

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Psychologie

Betreut von / Supervisor:

Mag. Dr. Ulrich Tran, Privatdoz.

Acknowledgments

Firstly, I would like to thank my thesis advisor Mag. Dr. Ulrich Tran, Privatdoz. for the continuous support throughout the whole master thesis. He has offered guidance in times of little progress and provided persistent help whenever needed. His encouraging and professional attitude made it a joy to work on the project.

Further, I would like to express appreciation to Assoz. Prof. MMag. DDDr. Martin Voracek, Privatdoz. for the additional input he provided when I first presented the preliminary results. His valuable suggestions laid the foundation for the final interpretation of the results which led to a major improvement of this thesis.

My special thanks go to Mag. Michael Kossmeier for his suggestions regarding the presentation of the results, and Mag. Danièle Beckius for proofreading the final manuscript.

Lastly, I would like to thank my family, friends, and colleagues. All of them contributed to the completion of this master thesis in various ways. This accomplishment would not have been possible without them.

Abstract

Socially desirable responding is defined as a subset of response biases in which the participants of a study do not produce accurate self-assessments, but influence their self-reports in a socially desirable fashion. The amount of bias is mainly associated with three factors: the research context, the questionnaire, and interindividual differences. The goal of the present master thesis was to evaluate the effects of socially desirable responding on the Big Five personality traits. The study was designed as a randomized controlled trial in which the participants ($N = 582$) were administered the same questionnaire at two measurement occasions. The control group completed the questionnaire without specific instructions at both measurement occasions, whereas the experimental group received an instruction at the second measurement occasion which was hypothesized to evoke socially desirable responding. Experimental data was analysed using structural equation models and socially desirable responding was modelled by common method variance factors. The results showed that mean differences that were caused by the experimental manipulation could be almost entirely modelled by means of one common method variance factor at the second measurement occasion that was identified as a factor of socially desirable responding. Covariances between personality traits were also reduced when accounting for this factor. In order to discover potential subgroups of responders, finite mixture models with subsequently increasing number of classes were fit to the data of the first measurement occasion. The resulting 2 classes differed with respect to Emotional Stability, Agreeableness, and Conscientiousness and social desirability was identified as a predictor of latent class membership. The results largely confirmed previous empirical findings, showing contextual effects and interindividual differences in socially desirable responding. Further, it is argued that the existence of a General Factor of Personality as a substantial personality trait is unlikely given the results that were obtained.

keywords social desirability, structural equation modelling, finite mixture modelling, personality;

Zusammenfassung

Sozial erwünschtes Antwortverhalten kann als Spezialfall von Antwortverzerrungen beschrieben werden, bei denen Personen in einer systematischen Weise sozial erwünschte Antworten in Fragebogenuntersuchungen geben, die mit der eigentlichen Selbsteinschätzung nicht übereinstimmen. Das Ausmaß des auftretenden sozial erwünschten Antwortverhaltens ist dabei maßgeblich von drei Faktoren beeinflusst: der Untersuchungssituation, dem Untersuchungsinstrument und interindividuellen Unterschieden. Die vorliegende Masterarbeit befasste sich mit dem Einfluss von sozial erwünschtem Antwortverhalten auf das Big Five Persönlichkeitsmodell. Die Probanden ($N = 582$) nahmen an einem randomized controlled trial mit zwei Messzeitpunkten teil, zu denen jeweils idente Fragebogenbatterien vorgelegt wurden. Die Kontrollgruppe bearbeitete die Fragebögen beide Male unter gleichen Bedingungen, während die Experimentalgruppe zum zweiten Messzeitpunkt eine Instruktion erhielt, die darauf ausgelegt war, sozial erwünschtes Antwortverhalten zu evozieren. Die Auswertung der Experimentaldaten erfolgte mittels Strukturgleichungsmodellen, in denen sozial erwünschtes Antwortverhalten als Methodenvarianzfaktor modelliert wurde. Es konnte gezeigt werden, dass die durch die Fragebogeninstruktion hervorgerufenen Mittelwertsunterschiede fast gänzlich durch Mittelwertsunterschiede in den Methodenvarianzfaktoren erklärt werden konnte. Die zuvor bestehenden Korrelationen zwischen Persönlichkeitsdimensionen konnten ebenfalls deutlich reduziert werden. Um interindividuelle Unterschiede in sozial erwünschtem Antwortverhalten zu testen, wurden Mischverteilungsmodelle an die Daten des ersten Messzeitpunktes angepasst. Die resultierenden zwei Klassen unterschieden sich hinsichtlich der Persönlichkeitsdimensionen Emotionale Stabilität, Verträglichkeit und Gewissenhaftigkeit. Soziale Erwünschtheit konnte als Prädiktor der Klassenzugehörigkeit identifiziert werden. Die Ergebnisse bestätigen vorhergehende Studien und zeigen den Einfluss von sozial erwünschtem Antwortverhalten auf Persönlichkeitserhebungen, insbesondere deren Kontextabhängigkeit und interindividuelle Unterschiede. Die aktuelle Masterarbeit trägt weiters zur Literatur des Generalfaktors der Persönlichkeit bei, dessen Existenz als substantieller Persönlichkeitsfaktor – gegeben den Ergebnissen – unwahrscheinlich erscheint.

Schlüsselwörter soziale Erwünschtheit, Strukturgleichungsmodellierung, finite Mischverteilungsmodellierung, Persönlichkeit;

Contents

1	Introduction	11
1.1	Defining the social desirability construct	11
1.2	Effects of social desirability on personality measures	12
1.2.1	Effects of questionnaire and context	12
1.2.2	Person-specific effects	14
1.3	Accounting for socially desirable responding	15
1.4	Goals of the master thesis	17
2	Materials & Methods	18
2.1	Participants	18
2.2	Procedure	19
2.2.1	Survey design	19
2.2.2	Data collection	20
2.3	Measurement Instruments	20
2.3.1	Personality	20
2.3.2	Social desirability	21
2.4	Analyses	22
2.4.1	Factorial analysis of the NEO-FFI and construction of a short form . .	22
2.4.2	Assessment of social desirability via Ziegler & Buehner's (2009) pro- cedure	22
2.4.3	Finite mixture modelling of the data	25
3	Results	27
3.1	Derivation of a short form NEO-FFI questionnaire	27
3.2	Structural equation modelling of the experimental data	28
3.2.1	Effects of socially desirable responding	28
3.2.2	Identification of faked personality traits	29
3.2.3	Modeling socially desirable responding	29
3.3	Finite mixture modelling of socially desirable responding	32
3.3.1	Model selection and goodness-of-fit	32
3.3.2	Class comparisons	33

4 Discussion	34
4.1 Effects of questionnaire and context	34
4.2 Person-specific effects	36
4.3 Limitations	37
4.4 Conclusion	38
References	39
List of Figures	47
List of Tables	49
Appendix	51
A Derivation of an item selection criterion	51
B Supplementary tables	52
C Survey information for data collection	55
D Survey instructions for participants	59
E Curriculum vitae	62

1 Introduction

Assessment and measurement of psychological constructs such as intelligence or personality are of utmost importance in psychological research. Oftentimes assessments are carried out in a fashion such that researchers are reliant on the self-reports of participants in a study. These self-report measurements are, however, possibly associated with systematic measurement error or response bias. Especially noncognitive assessments like personality questionnaires are criticised for being prone to response biases including extreme responding, acquiescence bias, or socially desirable responding (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Still, noncognitive assessments are utilized in a wide range of applications and play an important role in the prediction of career success (Ng, Eby, Sorensen, & Feldman, 2005) and workplace performance (Higgins, Peterson, Pihl, & Lee, 2007). Mainly for this reason much research is concerned with the influence of response bias on personality measurements, particularly socially desirable responding, which was shown to have great impact on self-reported personality profiles (Viswesvaran & Ones, 1999).

1.1 Defining the social desirability construct

Social desirability (socially desirable responding, social desirability bias, socially desirable response bias) is defined as a subset of response biases in which the participants of a study do not produce accurate self-assessments, but influence their self-reports in a socially desirable fashion (Hossiep, 2016). Many operationalizations of social desirability exist, ranging from simple one-dimensional scale constructions (e.g., Crowne & Marlowe, 1960; Edwards, 1957; Stöber, 1999) to two-level hierarchical factor structures (Paulhus, 2002). In general one can distinguish between *Self-Deception* and *Impression Management* (see e.g., Paulhus, 1984). *Self-Deception* refers to the tendencies of responding in a socially desirable manner with no intention to deliberately fake answers. Rather, socially desirable answers are given unconsciously and are assumed to have self-deceptive tendencies. *Impression Management* (also termed *Other-Deception*; Sackheim & Gur, 1978) includes deliberate lying or faking, and the "habitual presentation of a specific positive public impression" (Paulhus, 2002, p. 56). This leads to effects dependent on the assessment context, as there are context-specific differences in what is seen as a positive public impression (see e.g., Holden, 2007). It could also be shown that the amount of socially desirable responding varies between individuals

depending on their perception of the social desirability of the specific trait (Ludeke, Weisberg, & Deyoung, 2013). It follows logically that socially desirable response bias presents itself differently depending on the traits studied. Hence, the amount of socially desirable responding is also affected by the questionnaire used.

1.2 Effects of social desirability on personality measures

1.2.1 Effects of questionnaire and context

The Big Five were originally constructed as uncorrelated personality factors by the use of orthogonal rotation methods (varimax) of an exploratory factor analysis result (McCrae & Costa, 1987). Therefore the Big Five were assumed to be irreducible to higher dimensional factors. Research has shown, however, that the Big Five tend to be substantially correlated despite their theoretical independence (Van der Linden, te Nijenhuis, & Bakker, 2010) and subsequently potential higher order factor structures were investigated. This led to the development of hierarchical personality models such as the Alpha and Beta model (Digman, 1997), the Stability and Plasticity model (DeYoung, Peterson, & Higgins, 2002), and ultimately the General Factor of Personality (Musek, 2007). Questions have been raised whether or not higher level factors in these hierarchical models, particularly in the General Factor model by Musek (2007), reflect substantial personality traits (see e.g., Rushton & Irwing, 2008) or can be seen as methodological artefacts (e.g., socially desirable response bias; Davies, Connelly, Ones, & Birkland, 2015; Dunkel, van der Linden, Brown, & Mathes, 2016).

Studies in the context of social desirability do indeed show that socially desirable responding influences the covariance structure of questionnaires based on the Big Five personality traits under specific circumstances (e.g., Paulhus, Bruce, & Trapnell, 1995). Ziegler and Buehner (2009) illustrated that the revised *NEO Personality Inventory* (NEO-PI-R; Costa & McCrae, 1992) was subject to changes in factor correlations under an experimental manipulation. They found that participants – after being instructed to respond to the NEO-PI-R in a socially desirable manner – responded similarly to the scales independent of the personality traits. The similar response patterns in turn led to increased correlations between some, but not all personality traits when compared to a control group. A study by Pauls and Crost (2005) further underpinned these findings, but also showed context-dependent

effects on the correlational structure of the NEO-PI-R. Under general faking instructions¹ all NEO-PI-R scales were affected, whereas only few trait intercorrelations increased when participants were instructed to implicitly respond in a socially desirable way by employing applicant instructions. Which correlations increased was solely dependent on the job description given in the questionnaire instructions (Pauls & Crost, 2005, Table 2, p. 303). Both findings are consistent with the operationalization of social desirability (Impression Management) by Paulhus (2002) and the assumption that the amount of socially desirable responding is largely dependent on the context of the assessment (see Section 1.1). They also indicate that a socially desirable response bias might be the underlying cause of higher level factors rather than substantial personality traits.

Empirical evidence is still inconclusive as effects of social desirability on the questionnaires covariance structure could not be found consistently in all studies (see e.g., Hogan, Barrett, & Hogan, 2007). Ziegler and Buehner (2009) argued that inconsistently detected effects could be due to sample heterogeneity in application settings and/or the use of social desirability scales to detect socially desirable responding. Firstly, sample heterogeneity is brought forward as a confounding factor for assessments of social desirability and a possible source for contradictory findings. This argument is based on the assumption that socially desirable responding differs based on which traits are thought to be socially desirable. Indeed it could be shown that social desirability of specific traits do vary between job positions (manager, employee) and professions in applicant settings (Birkeland, Manson, Kisamore, Brannick, & Smith, 2006). If the study design does not account for those confounders (e.g., by pooling multiple job positions or professions in one sample) the effects of social desirability on specific personality traits might be underestimated. Secondly, the issues of social desirability scales must be addressed. Their use for detection of socially desirable responding is problematic mainly because of two reasons: (1) they show considerable correlations with personality traits (Ones, Viswesvaran, & Reiss, 1996), even in low stakes (non-applicant) settings (Musch, Brockhaus, & Bröder, 2002; Stöber, 2001), and (2) their psychometric properties are insufficient (Li & Bagger, 2007; Tran, Stieger, & Voracek, 2012). For those reasons it is often advised against the use of social desirability scales (see e.g., Cervellione, Lee, & Bonanno, 2009). Overall it seems plausible that effects of social desirability bias on the covariance structure of personality questionnaires exist and can be

¹"Present yourself as favourably as possible" (Pauls & Crost, 2005, p. 300)

found if the confounding factors are considered.

Another well established finding is that socially desirable responding affects trait questionnaire mean scores in both between- and within-subjects studies (Birkeland et al., 2006; Viswesvaran & Ones, 1999). Depending on the research context, changes in mean scores differ for the Big Five personality traits. In job application settings socially desirable responding is typically more pronounced in traits which are perceived to be more relevant by the participants. The effects are consistently largest for Neuroticism (Emotional Stability) and Conscientiousness, whereas only small effects occur for Extraversion, Agreeableness, and Openness in experimental studies which use instructions to mimic application settings (Birkeland et al., 2006; Ziegler & Buehner, 2009).

1.2.2 Person-specific effects

Although much research has gone into socially desirable responding, emphasis has mostly been placed on the effects of socially desirable responding on the questionnaire and context level and person-specific effects remained largely ignored. With the relatively recent emergence of person-oriented statistical methods, however, the investigation of person- and subpopulation-specific effects found its way into assessment of personality (Maij-de Meij, Kelderman, & van der Flier, 2008) and social desirability research (Leite & Cooper, 2010; Ziegler, Maaß, Griffith, & Gammon, 2015).

Despite the availability of appropriate modelling techniques such as latent classification methods only few studies to date examined the effects of social desirability on personality measures on the person and subpopulation level. Holden and Book (2009) showed that an Impression Management scale was able to discern between groups instructed to respond differently on personality measures. Evidence was provided that latent classification is suitable to explore differences in socially desirable responding in experimental studies. Another study by Maij-de Meij et al. (2008) investigated group-specific effects of socially desirable responding on Neuroticism and Extraversion in a personnel selection procedure. It was concluded that a model with three latent classes fit the data best and classes differed with respect to social desirability scores. The effects were especially pronounced when modelling Neuroticism, but still present in the Extraversion model. These preliminary studies and methodological advancements have shown that population level analysis might be inappropriate when modelling socially desirable responding and latent classification should

be taken into consideration to increase prediction accuracy.

1.3 Accounting for socially desirable responding

Many approaches to model socially desirable responding have been developed (e.g., Ferrando, 2005; Lorenzo-Seva & Ferrando, 2012; Saar, Aavik, & Konstabel, 2012; Ziegler & Buehner, 2009). Within these, methods based on repeated measurement need to be distinguished from other methods (for a review see also Burns & Christiansen, 2011). In applied research and personnel selection non-repeated measurement methods are preferred because of their ease of application, but they suffer from two serious drawbacks:

1. *Most methods rely on the usage of social desirability scales.*

As outlined above, social desirability scales do not meet commonly demanded psychometric standards (see Section 1.2.1). Relying on these scales as a measure of response bias is problematic, because substantial correlations with personality traits do not allow for unambiguous interpretations of social desirability scores. High or low scores could be an indicator of extreme personalities instead of socially desirable responding. Still, social desirability scales are frequently used to correct mean scores in personality assessments (Goffin & Christiansen, 2003).

2. *Separation of trait variance and method variance is not possible.*

Methods based on the covariance structure of a questionnaire (e.g., Covariance Index) make use of the increased correlations between personality traits when socially desirable responding occurs (Burns & Christiansen, 2011). These methods rely on the idea that social desirability can be treated as systematic measurement error (common method variance; Podsakoff et al., 2003) which is shared by personality measures and therefore leads to high correlations between traits. However, a clear separation of common method variance and trait variance is not possible in a single measurement design due to the fact that increased correlations can also be a result of other forms of measurement error (e.g., item characteristics) or substantial higher level traits.

In the context of psychometric evaluation of questionnaires repeated measurement methods are generally preferred, because they do not rely on social desirability scales and socially desirable responding can be evoked by experimental manipulation. A promising

modelling approach was proposed by Ziegler and Buehner (2009), combining ideas of common method variance (Podsakoff et al., 2003; Podsakoff, MacKenzie, & Podsakoff, 2012) and repeated measurement modelling using structural equations:

Two groups (control group, experimental group) are administered the same questionnaires at two measurement occasions. At the first measurement participants fill in the questionnaire regularly without specific instructions. At the second measurement occasion it is differentiated between the control group and the experimental group, which receive the same instruction or a specific instruction to respond socially desirable respectively. The data of both measurements is modelled regularly but also includes common method variance factors. The experimental procedure and testing of a set of hypotheses allows unambiguous interpretation of the common method variance factors as socially desirable response factors, because the only difference between groups is the social desirability instruction.² (Ziegler & Buehner, 2009)

The benefits of Ziegler and Buehner's (2009) approach is that both hypotheses about the effects of social desirability on the covariance structure, as well as on the mean scores can be addressed simultaneously. Assuming that Big Five personality factors are actually basic and irreducible to higher dimensional factors, no residual correlations between personality traits should remain when accounting for social desirability. If unexplained correlations between personality traits remain, it suggests that substantial higher level factors exist. Modelling the mean structure should theoretically also account for changes in the Big Five mean scores. It is expected that group differences in the latent personality factor means disappear and are instead represented as a latent mean difference in the common method variance factors.

² For a more detailed description of the models and hypotheses see Section 2.4.2.

1.4 Goals of the master thesis

In the previous sections it was established that social desirability bias has potentially big effects on Big Five personality measures. Effects on both the covariance structure of the questionnaire and personality trait mean scores are commonly found in the literature and can be modelled in different ways. The approach by Ziegler and Buehner (2009) stands out as it allows simultaneous modelling of the covariance and the mean structure. Additionally, a clear interpretation of the common method variance factors as socially desirable responding is provided. This warrants replication studies with other personality questionnaires and under different experimental manipulations in both applicant and non-applicant settings.

Application of person-oriented methods in social desirability and personality suggest that both experimental data and personnel selection data yield multiple groups which can be distinguished by their scores in social desirability measures (Holden & Book, 2009; Maij-de Meij et al., 2008). To the best of my knowledge studies have yet to investigate latent classifications in low-stakes and non-experimental settings. Further, only selected personality traits were investigated in the past and no results on joint Big Five personality profiles are available.

The proposed goal of the master thesis at hand was to apply the experimental procedure derived by Ziegler and Buehner (2009) to study social desirability in a non-applicant setting, and expand person-oriented research in the domain of personality by investigating possibly naturally occurring subgroups which differ in their socially desirable response tendencies.

2 Materials & Methods

2.1 Participants

The sample was comprised of $N = 582$ participants, who were assigned to either one the experimental conditions via block-randomization to ensure approximately equal group sizes (experimental group: $n = 292$, control group: $n = 290$). Mean age at the time of investigation was 34.67 years ($SD = 13.93$) and did not differ across groups (experimental group: $M = 34.54$, $SD = 13.71$; control group: $M = 34.79$, $SD = 14.18$, $t(575) = 0.21$, $p = .833$). As shown in Figure 1 the sample was not representative of the general population regarding the age distribution, but can be considered a community sample as a result of the data acquisition procedure (see Section 2.2.2).

Participant sex was distributed approximately even, with 51.7% of the participants being female (experimental group: 52.7%, control group: 50.7%). There were only marginal deviations from uniform distribution, $\chi^2(3) = .946$, $p = .815$.

Participants also described themselves as highly educated in comparison to the general population in Austria (Statistik Austria, 2015). The median highest qualification was the *general qualification for university entrance (Matura)*, again with no differences between groups (see Figure 2 for a more detailed description of the educational background).

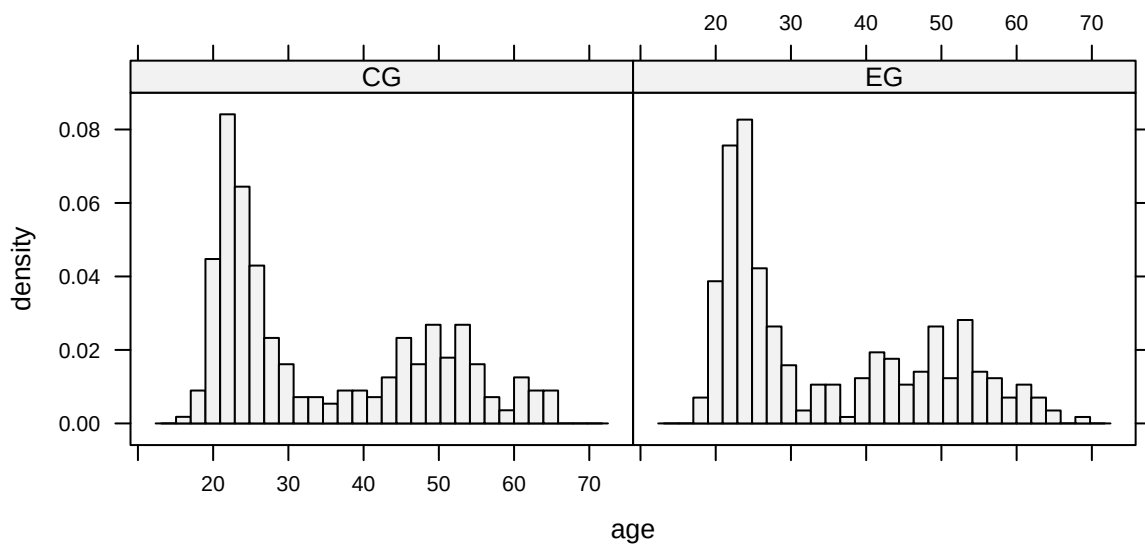


Figure 1. Age distribution in the sample (by group).

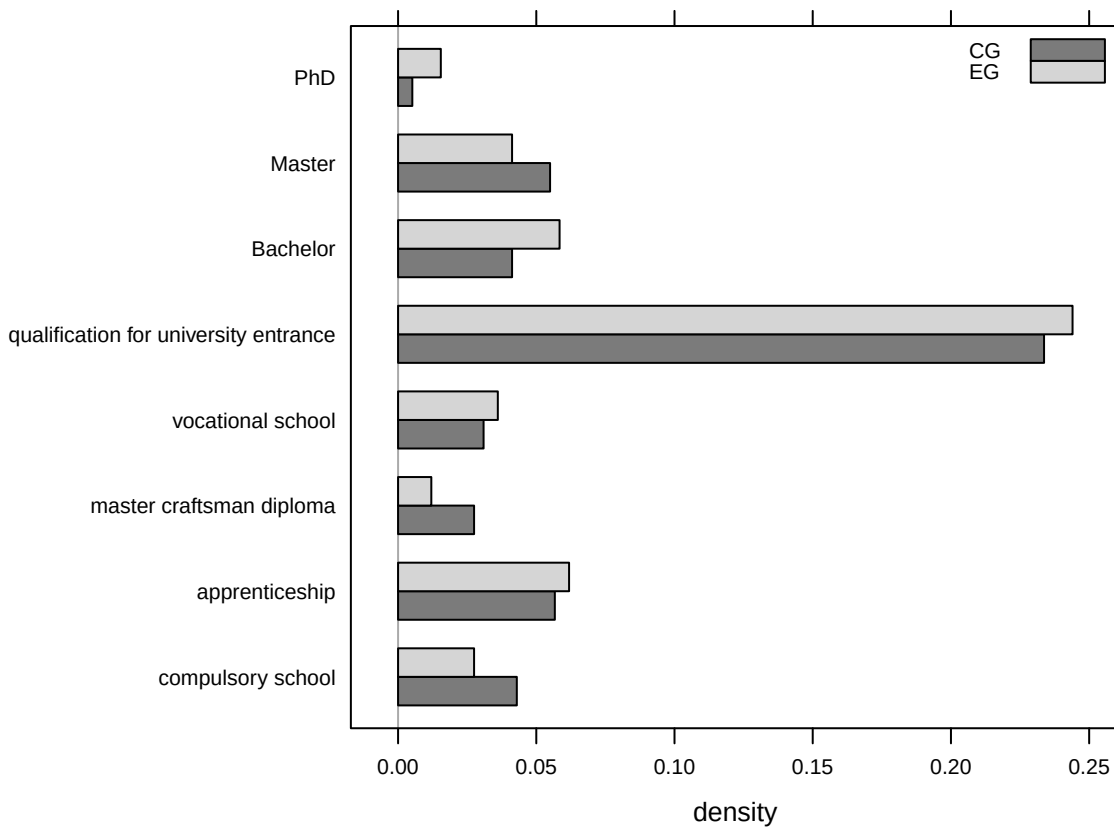


Figure 2. Educational background in the sample (by group).

2.2 Procedure

2.2.1 Survey design

The study was designed as a randomized controlled trial with two measurement occasions. Participants were assigned to either the experimental group or the control group and asked to fill in a collection of questionnaires as well as socio-demographic data. Random assignment was guaranteed by the instructions given to the students who collected the data (see Section 2.2.2 for a short description). The questionnaires used were the (1) *NEO Five Factor Inventory* (Borkenau & Ostendorf, 1993; Costa & McCrae, 1989), (2) *Fragebogen zur Geschlechtstypizität* (Altstötter-Gleich, 2004), (3) *Positive and Negative Affect Schedule* (Krohne, Egloff, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen, 1988), (4) *Behavioral Inhibition System/Behavioral Approach System Scales* (Carver & White, 1994; Strobels, Beauducel, Debener, & Brocke, 2001), (5) *Balanced Inventory of Desirable Responding* (Musch et al., 2002; Paulhus, 1994), and (6) *Stöber's Social Desirability Scale* (Stöber, 1999). At the first measurement occasion both groups completed the questionnaires without

any specific instruction, whereas group specific instructions were introduced for the second measurement. The control group again completed the questionnaires without any specific instructions and the experimental group was told to "*consider this time how a person who is admired and praised by others would answer the questions.*" [Überlegen Sie sich dieses Mal jedoch, wie eine Person die Fragen beantworten würde, die von anderen bewundert und gelobt wird.], which was hypothesized to evoke socially desirable responding in a general sense.

2.2.2 Data collection

The data were acquired by undergraduate students over the course of the summer term 2013 as a part of their bachelor thesis course (Bachelorarbeit I) at the University of Vienna. The students were instructed on the overall subject matter, survey design and the measurement instruments used (please refer to Appendix C for the full instructions received by the students). They were specifically briefed on the block randomization process and not to go beyond 7 days between measurements. The mean time difference between measurement occasions was 3.97 days ($SD = 1.99$ days) and data for only 22 participants (3.78%) exceeded the criterion of 7 days, with a maximum of 17 days.

2.3 Measurement Instruments

2.3.1 Personality

NEO Five Factor Inventory. Assessment of the Big Five was carried out using the *NEO Five Factor Inventory* (NEO-FFI; Costa & McCrae, 1989) in its German translation by Borkenau and Ostendorf (1993). The NEO-FFI consists of five theoretically uncorrelated scales: *Neuroticism*, *Extraversion*, *Openness to Experience*, *Agreeableness*, and *Conscientiousness*. Each scale is measured by 12 items which results in a total amount of 60 items for the whole questionnaire. The answers are given on a 5-point Likert scale ranging from "-" (strong disagreement) to "++" (strong agreement). All items have been recoded according to the manual (Borkenau & Ostendorf, 1993) and the Neuroticism scale has been inverted to depict a measure of *Emotional Stability* in all further analyses. Internal consistency (Cronbach's α) ranged from .75 for Agreeableness to .86 for Conscientiousness indicating adequate reliability of the scales. The mean interitem correlations spanned between .21 (Extraversion, Agreeableness) and .35 (Conscientiousness). All mean item-total correlations were greater

than .50.

Lack of factorial validity in the full 60-item questionnaire version led to the construction of a 30-item short form (see Section 2.4.1 and 3.1). For an overview of reliability measures of both questionnaire forms see Table 1 (Section 3.1, p. 27).

2.3.2 Social desirability

Balanced Inventory of Desirable Responding. The *Balanced Inventory of Desirable Responding* (BIDR; Paulhus, 1994) is a measurement instrument based on a two-factor conceptualisation of socially desirable responding (see e.g., Paulhus, 2002) and differentiates between *Self-Deceptive Enhancement* (SDE; e.g., "I am a completely rational person." [*Ich bin ein vollkommen rational denkender Mensch.*]) and *Impression Management* (IM; e.g., "I never swear." [*Ich fluche niemals.*]). The scales consist of 10 items each. An alternative answer format on a 7-point Likert-type scale (1 = *total disagreement* – 7 = *total agreement*) was chosen as it was shown to be preferable to dichotomous scoring in terms of internal consistency and convergent validity (Stöber, Dette, & Musch, 2002). In the current study, the German version of the questionnaire was used (Musch et al., 2002). Scale scores for each person were computed as the sum of item scores of the respective scale after negatively keyed items were reversed. Internal consistency was comparable to previously published results for the BIDR (Li & Bagger, 2007), $\alpha = .64$ for the Self-Deceptive Enhancement scale and $\alpha = .70$ for Impression Management. Low mean inter-item correlations for both Self-Deceptive Enhancement, $M(r_{ij}) = .14$, and Impression Management, $M(r_{ij}) = .20$, revealed the heterogeneity of measurement for the social desirability construct within scales.

Social Desirability Scale. Stöber's *Social Desirability Scale* is a social desirability measurement originally developed in German (Stöber, 1999). It approaches social desirability bias as a one-dimensional measurement and contains 17 items (e.g., "There has been at least one occasion when I failed to return an item that I borrowed." [*Ich habe schon einmal geliehene Sachen nicht zurückgegeben.*], reverse keyed). Instead of the original scoring with a dichotomous answer format, a 7-point Likert scale (1 = *total disagreement* – 7 = *total agreement*) was used as suggested by Tran et al. (2012). Cronbach's α was .75. The mean inter-item correlation was low, $M(r_{ij}) = .16$, suggesting that the measure is not one-dimensional.

2.4 Analyses

2.4.1 Factorial analysis of the NEO-FFI and construction of a short form

Factorial validity of the NEO-FFI was assessed by fitting multiple *structural equation models* (SEM) to the data of the first measurement occasion. Both orthogonal and correlated five-factor models were considered.

Following the SEM analyses, a short form of the questionnaire was constructed based on the results of *exploratory structural equation modeling* (ESEM; Asparouhov & Muthén, 2009). ESEM incorporates approaches from both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) in that they introduce free estimation of factor loadings (EFA) on a predetermined factor structure (CFA) and also provide goodness-of-fit measures from standard SEM analysis. Therefore ESEM specifically allows investigation of item performance based on their factor loadings. For the construction of the short form items were selected such that the loadings on their designated factors were high, and cross-loadings on other factors were low. An *item selection criterion* (ISC) was introduced to allow comparisons of item performance (for a full derivation of the ISC please refer to Appendix A). For each factor the 6 items with the highest ISC were selected, that is items with the highest ratio of (squared) loadings on the designated factor to mean (squared) cross loadings. Factorial validity of the short form questionnaire was then evaluated using standard confirmatory factor analysis. Both orthogonal and oblique factor solutions were checked and compared.

All models were fit using the mean and variance adjusted least squares estimator (WLSMV) in MPLUS 7 (Muthén & Muthén, 1998-2010). The DIFFTEST option was used for model comparisons (Asparouhov & Muthén, 2006). Assessment of model fit was the same for all SEM analyses and based on benchmarks by Hu and Bentler (1999). Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and root mean square error of approximation (RMSEA) including its 90% confidence interval are reported.

2.4.2 Assessment of social desirability via Ziegler & Buehner's (2009) procedure

Testing of the set of hypotheses derived by Ziegler and Buehner (2009) was carried out in five steps. In a first step descriptive statistics (correlations, means, and variances) of the personality scales were calculated for both measurement occasions, which showed the impact of the questionnaire instructions on the experimental group and serve as a reference

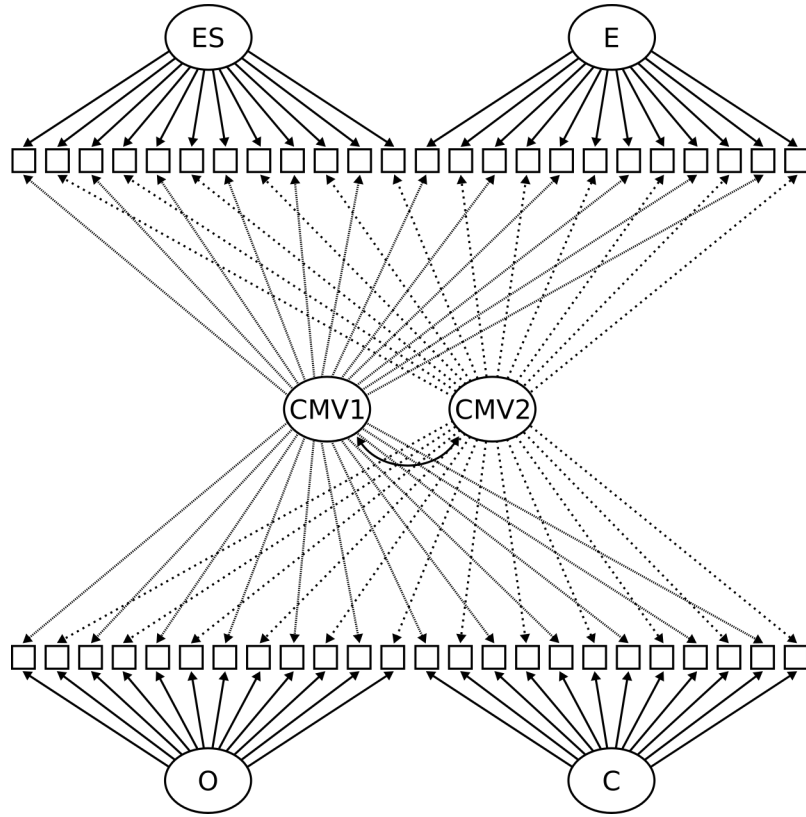


Figure 3. Full structural equation model controlling for social desirable responding.

Note. ES = Emotional Stability, E = Extraversion, O = Openness, C = Conscientiousness, CMV1 = common method variance (first measurement), CMV2 = common method variance (second measurement); item residuals, correlations of item residuals, and personality factor correlations have been omitted to facilitate graphical representation.

for further statistical analysis.

Secondly, the identification of scales affected by socially desirable responding was conducted using analysis of covariance (ANCOVA) for each scale. The outcome, personality score at the second measurement occasion, was predicted by group membership, while personality score at first measurement occasion was included as a metric covariate. In this case the ANCOVA approach is preferable to repeated measure analysis of variance (ANOVA), because the effect of interest is modelled as the main effect of the experimental condition variable rather than the interaction effect *experimental condition* $\times$ *measurement occasion*. This procedure leads to increased power of the analysis compared to more traditional approaches such as repeated measures ANOVA (Vickers & Altman, 2001). Estimates of population effect sizes (Hedges *g*; Cooper, Hedges, & Valentine, 2009) and their 95% confidence intervals were calculated and reported for the group differences.

After identification of faked personality traits structural equation models were fit to the data in a third step of the analysis. The SEM contained the data for all Big Five person-

ality traits that showed socially desirable responding. Each personality trait was modelled as a single latent variable measured by the corresponding items of both time points (for an overview of the model see Figure 3). To account for the dependence of the repeated measurement within subjects, the error terms were allowed to correlate between identical items across measurement occasions. Additionally two correlated latent common method variance factors were introduced to the model, which were measured by all items of the first or second measurement occasion respectively. Loadings on the personality factors and item thresholds were constrained to be invariant across groups to allow estimation of a mean structure. In the control group means were set to zero for all but the common method variance factor at the first measurement occasion, whereas means in the experimental group were estimated freely. The common method variance factor at the first measurement occasion was constrained to be equal across groups. A full graphical representation for a single personality trait is presented in Figure 4. To account for the ordinal nature of the NEO-FFI items, the SEM was estimated using the WLSMV estimator in MPLUS 7 (Muthén & Muthén, 1998-2010). Goodness of fit was assessed by CFI, TLI, and RMSEA (90% CI). Models with correlated and uncorrelated latent personality factors were compared with the DIFFTEST option (Asparouhov & Muthén, 2006).

The next step of the analysis included assessment of the covariance and mean structure of the model. Standardized factor covariances were compared between groups with the hypothesis that all correlations between personality traits should be low when accounting for socially desirable responding. Latent mean differences between groups and their effect sizes were also investigated. It was expected that differences between groups in the personality traits disappear. If socially desirable responding is the only source of the mean differences, it should be wholly represented in the mean difference of the common method variance factor at the second measurement occasion and a higher mean score should occur for the experimental group.

To further ensure that the interpretation of the common method variance factors as socially desirable responding is acceptable, correlations within groups were compared. According to Ziegler and Buehner (2009) common method variance factors should correlate highly in the control group, but not in the experimental group. A low correlation in the experimental group reflects the change in situational demand associated with the social desirability instruction, which is not present in the control group.

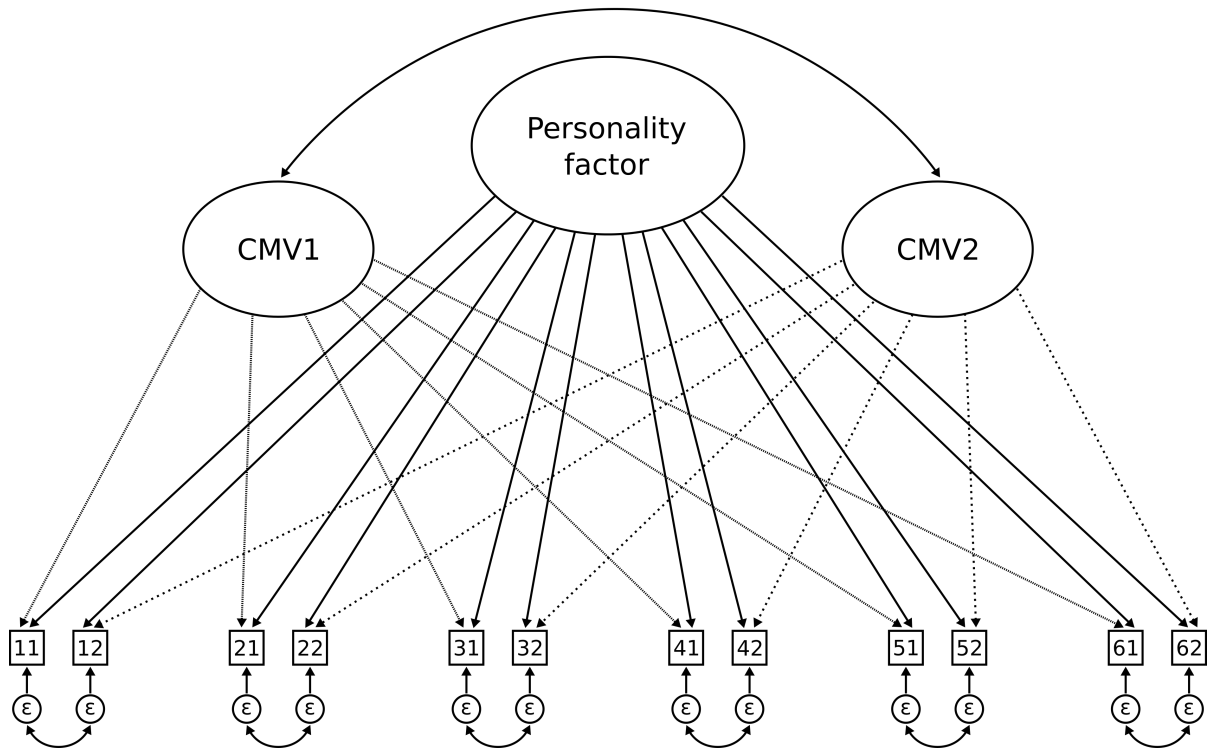


Figure 4. Detail of the SEM illustrating a single personality trait.

Note. The first digit indicates item number and the second digit shows measurement occasion (e.g., 41: 4th item, first measurement occasion).

Additionally – to address convergent validity – correlations were computed for SDS-17 mean scores, BIDR Self-Deceptive Enhancement scores, and BIDR Impression Management scores with the common method variance factor scores of both groups. The lack of social desirability instruction in the control group suggested that correlations should be present predominantly for the SDS-17 and BIDR Self-Deceptive Enhancement scales. They were further assumed to be stable over time. Similar correlations were expected for the experimental group at the first, but not second measurement occasion. Here a change to deliberate socially desirable responding was expected, which should manifest in correlations with the BIDR Impression Management scale.

All analyses except the estimation of the SEM were performed in R (Version 3.3.0; R Core Team, 2016).

2.4.3 Finite mixture modelling of the data

In order to discover potential subgroups of responders, a finite mixture model was fit to the data of the first measurement occasion (see Figure 5). In the model a latent categorical factor (latent class factor) was measured by the mean score of all five NEO-FFI scales

(Emotional Stability, Extraversion, Openness, Agreeableness, and Conscientiousness). Additionally, it was desired to explore the effect of social desirability on latent class membership. Mean scores of the SDS-17, BIDR Impression Management, and BIDR Self-Deceptive Enhancement were therefore included as predictors of the latent class factor. Measurement models of the questionnaires were omitted from the model to decrease computation time as well as increase computational stability of the results.

The models were fit with subsequently increasing number of classes. To assess the most likely number of classes, an approach by Nylund, Asparouhov, and Muthén (2007) was used. First, candidate models were selected based on Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and sample size adjusted Bayesian Information Criterion (aBIC). The Likelihood-Ratio between two candidate models was subsequently compared using Bayes Factor and was also tested with a bootstrap Likelihood-Ratio Test statistic (BLRT; Asparouhov & Muthén, 2012; Nylund et al., 2007). The resulting classes were then compared in an exploratory fashion. Each participant was assigned a group membership based on their highest latent class probability, which were then evaluated with respect to additional measures such as participants age, educational level, and gender using appropriate (generalized) linear models.

Mixture models were fit in MPLUS 7 (Muthén & Muthén, 1998-2010) using maximum likelihood estimation with robust standard errors (MLR). BLRT statistics were requested via the TECH14 option (Asparouhov & Muthén, 2012). Post classification analyses and tests were performed in R (Version 3.3.0; R Core Team, 2016).

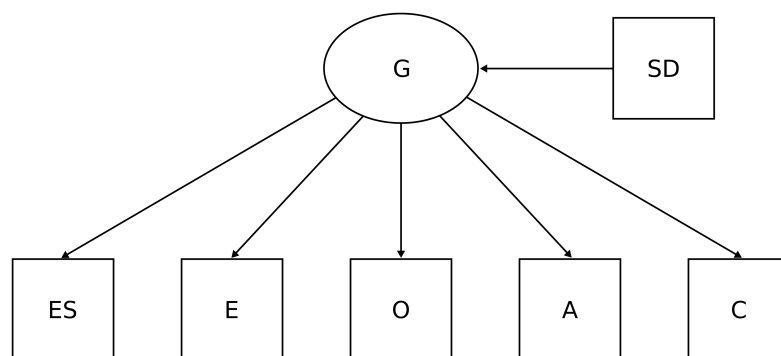


Figure 5. Finite mixture model diagram.

Note. G = latent class factor, SD = social desirability, ES = Emotional Stability, E = Extraversion, O = Openness, A = Agreeableness, C = Conscientiousness; correlations between personality traits have been omitted to facilitate graphical representation.

3 Results

3.1 Derivation of a short form NEO-FFI questionnaire

Factorial validity of the NEO-FFI was assessed using the data of the first measurement occasion. Goodness of fit measures revealed weak model fit of the five factor model, $\chi^2(3645) = 8114.62, p < .001$, CFI = .632, TLI = .642, RMSEA = .065 [.063, .067]. Allowing correlated factors improved the model fit significantly, $\Delta\chi^2(20) = 472.72, p < .001$. Despite the improvement, the model fit was still unsatisfactory, $\chi^2(3625) = 6324.86, p < .001$, CFI = .778, TLI = .783, RMSEA = .051 [.049, .053]. In order to reveal sources of misfit and poorly discriminating items an ESEM was fit to the data. The free estimation of factor loadings and factor covariances improved the goodness of fit considerably, $\chi^2(1480) = 2523.06, p < .001$, CFI = .924, TLI = .910, RMSEA = .035 [.032, .037], and also revealed weakly discriminating items and items with high cross loadings in the questionnaire (see Table 7 of ESEM factor loadings in Appendix B). Items for the short form were subsequently selected based on the previously introduced ISC (see Appendix A). CFA of the short form questionnaire revealed that the model with orthogonal personality factors again failed to describe the data well, $\chi^2(917) = 2265.65, p < .001$, CFI = .759, TLI = .772, RMSEA = .071 [.067, .075]. Noticeable improvements could be achieved by allowing factor intercorrelations, $\Delta\chi^2(20) = 227.25, p < .001$. The final model was therefore based on the short form questionnaire with correlated latent factors, which showed acceptable fit with respect to RMSEA as well as higher fit indices compared to the other models, $\chi^2(897) = 1736.90, p < .001$, CFI = .850, TLI = .855, RMSEA = .053 [.057, .061]. Descriptive measures and measures of reliability are reported in Table 1.

Table 1. Reliability measures of the full and short form NEO-FFI scales.

Scale	full form (60 items)				short form (30 items)			
	<i>M</i> (<i>SD</i>)	α	$M(r_{ij})$	$M(r_{it})$	<i>M</i> (<i>SD</i>)	α	$M(r_{ij})$	$M(r_{it})$
Emotional Stability	2.40 (0.67)	.84	.31	.61	2.38 (0.76)	.76	.35	.68
Extraversion	2.45 (0.52)	.76	.21	.53	2.36 (0.63)	.70	.28	.63
Openness	2.44 (0.61)	.78	.22	.54	2.38 (0.82)	.78	.37	.69
Agreeableness	2.68 (0.50)	.75	.21	.52	2.70 (0.64)	.70	.28	.63
Conscientiousness	2.69 (0.64)	.86	.35	.63	2.72 (0.67)	.77	.36	.68

Note. α = Cronbach's α ; $M(r_{ij})$ = mean interitem correlation; $M(r_{it})$ = mean item-total correlation. Note also that Cronbach's α is dependent on questionnaire length and therefore expected to be lower in the short form of the questionnaire (Sijtsma, 2008).

Table 2. Means and standard deviations for all variables.

	T1		T2	
	CG	EG	CG	EG
NEO-FFI				
Emotional Stability	2.39 (0.75)	2.38 (0.77)	2.44 (0.75)	3.00 (0.68)
Extraversion	2.37 (0.61)	2.35 (0.65)	2.37 (0.62)	3.14 (0.61)
Openness	2.40 (0.81)	2.36 (0.82)	2.39 (0.85)	2.54 (0.80)
Agreeableness	2.73 (0.61)	2.67 (0.68)	2.69 (0.63)	2.69 (0.83)
Conscientiousness	2.71 (0.67)	2.73 (0.67)	2.67 (0.67)	3.27 (0.56)
BIDR				
Self-deceptive enhancement	42.61 (6.60)	41.54 (8.02)	42.79 (6.67)	50.31 (10.26)
Impression management	38.58 (9.82)	37.68 (10.51)	39.79 (11.13)	44.34 (13.94)
SDS-17	4.85 (0.71)	4.81 (0.77)	4.93 (0.75)	5.42 (1.10)

Note. T1 = first measurement occasion; T2 = second measurement occasion; CG = control group; EG = experimental group;

3.2 Structural equation modelling of the experimental data

3.2.1 Effects of socially desirable responding

Descriptive statistics for all measures in use is displayed in Table 2. It is apparent that means as well as standard deviations of personality traits remained mostly unchanged across measurements in the control group, whereas increased scores in Emotional Stability, Extraversion, Openness, and Conscientiousness can be observed in the experimental group.

Correlations between personality traits at the first measurement occasions were small to moderate, ranging from .00 to .37, and did not differ between groups (see Table 3, column T1). At the second measurement occasion correlations in the control group remained almost unchanged (range: .00 – .36). The experimental group exhibited increased correlations throughout (range: .34 – .51; see Table 3, column T2). Difference in correlation coefficients across measurement occasions³ ranged from –.01 (Emotional Stability with Extraversion) to .14 (Emotional Stability with Conscientiousness) in the control group and from .12 (Agreeableness with Extraversion) to .48 (Openness with Conscientiousness) in the experimental group. The groups are almost completely separated by the median difference in correlation coefficients, $Md = .13$ (see Figure 6), which indicates that the change in correlation coefficients is caused by the social desirability instruction received by the experimental group at the second measurement occasion.

³ $\Delta r = r_{T2} - r_{T1}$

Table 3. NEO-FFI correlations with 95% confidence intervals (raw data).

Correlation	T1		T2	
	CG	EG	CG	EG
Emotional Stability with				
Extraversion	.37 [.26, .46]	.24 [.13, .34]	.36 [.25, .45]	.51 [.42, .59]
Openness	-.04 [-.15, .08]	-.04 [-.15, .07]	.00 [-.11, .12]	.42 [.32, .51]
Agreeableness	.08 [-.04, .19]	.14 [.03, .25]	.20 [.09, .31]	.34 [.23, .43]
Conscientiousness	.22 [.11, .33]	.30 [.19, .40]	.36 [.26, .46]	.51 [.42, .59]
Extraversion with				
Openness	.14 [.03, .25]	.17 [.06, .28]	.21 [.10, .32]	.44 [.34, .53]
Agreeableness	.25 [.14, .35]	.27 [.16, .38]	.32 [.21, .42]	.39 [.29, .49]
Conscientiousness	.22 [.11, .33]	.22 [.11, .33]	.25 [.14, .36]	.49 [.40, .57]
Openness with				
Agreeableness	.18 [.07, .29]	.08 [-.03, .19]	.25 [.14, .36]	.43 [.33, .52]
Conscientiousness	-.20 [-.31, -.08]	-.12 [-.23, -.01]	-.19 [-.30, -.08]	.36 [.25, .45]
Agreeableness with				
Conscientiousness	.20 [.09, .31]	.24 [.13, .35]	.25 [.13, .35]	.43 [.33, .52]

Note. T1 = first measurement occasion; T2 = second measurement occasion; CG = control group; EG = experimental group;

3.2.2 Identification of faked personality traits

ANCOVAs revealed that four out of five personality traits were affected by socially desirable responding. Estimates of population effect sizes indicate strong effects of the experimental condition for Extraversion, $F(1, 579) = 305.62, p < .001, g = 1.33 [1.16, 1.49]$, and Conscientiousness, $F(1, 579) = 195.10, p < .001, g = 0.99 [0.84, 1.14]$, at least a moderately sized effect for Emotional Stability, $F(1, 579) = 118.02, p < .001, g = 0.81 [0.65, 0.96]$, and a small effect for Openness, $F(1, 579) = 13.45, p < .001, g = 0.23 [0.10, 0.35]$. Only Agreeableness mean scores appeared to be unaffected by the socially desirable responding instruction, $F(1, 579) = 0.30, p = .586, g = 0.04 [-0.11, 0.19]$.

3.2.3 Modeling socially desirable responding

Based on the prior ANCOVA analyses only Emotional Stability, Extraversion, Openness, and Conscientiousness were included in the SEM. Goodness of fit measures indicate acceptable to good fit of the model, $\chi^2(2202) = 3235.66, p < .001, CFI = .946, TLI = .944, RMSEA = .040 [.037, .043]$. The model with orthogonal personality factors fit the data worse, $\Delta\chi^2(12) = 70.93, p < .001$, and was therefore not considered in further analyses.

Overall, point estimates of the standardized factor covariances between personality traits in the oblique factor model were small and the pattern was similar for both groups (see Ta-

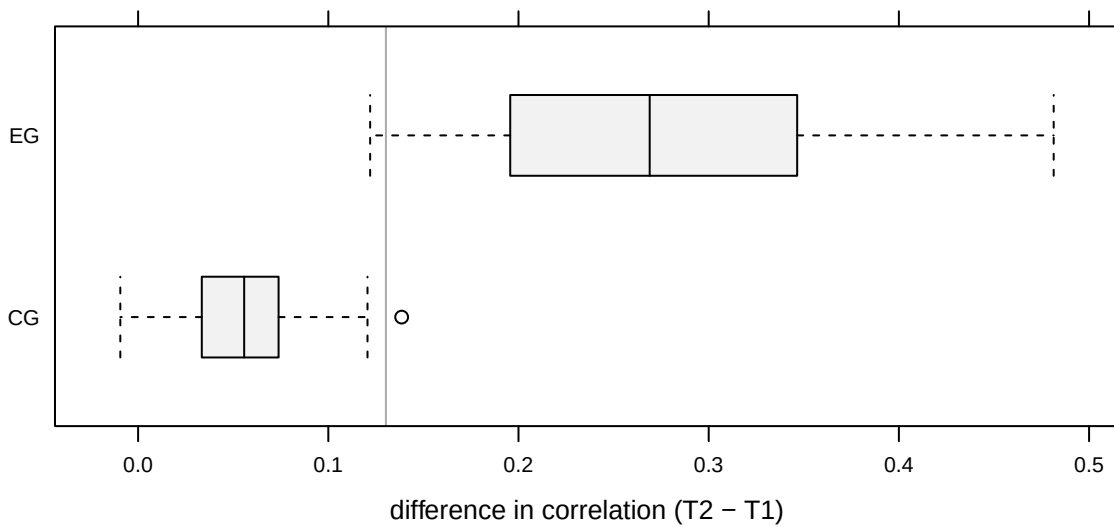


Figure 6. Boxplot of difference in correlation coefficients across measurement occasions.
Note. The gray vertical line indicates the median difference in correlation for the whole data set.

ble 4). In both groups substantial correlations between Extraversion and Openness (CG: $r = .27$, EG: $r = .31$), and Extraversion and Conscientiousness (CG: $r = .19$, EG: $r = .22$) remained. Additionally, in the control group only, correlations between Extraversion and Emotional Stability, $r = .18$, and Openness and Conscientiousness, $r = -.15$, were observed. After comparing the correlations of the raw data with the correlations estimated by the SEM, it was concluded that the intercorrelatedness of the personality traits was reduced when accounting for socially desirable responding – especially in the experimental group.

On the level of factor loadings it was observed that all items still significantly loaded on their intended personality factor (see Table 8 in Appendix B). Loadings on the common method variance factor at the first measurement occasion were constrained to equality across groups and ranged between $-.20$ (Item 13) and $.62$ (Item 6). All items of the personality factors Emotional Stability and Conscientiousness shared common method variance, whereas 3 items of Extraversion, and 2 items of the Openness scale did not. The sign of factor loadings on the common method variance factors was positive for all items except items assigned to the Extraversion scale, which showed a negative sign. The loading pattern on the common method variance factor at the second measurement occasion was virtually identical in the control group, but factor loadings were higher in the experimental group (range: $.21 - .74$). These results, together with the reduced correlations between personality factors, are evidence for the increased impact of socially desirable responding on the

Table 4. NEO-FFI correlations with 95% confidence intervals (SEM).

Correlation	CG	EG
Emotional Stability with		
Extraversion	.18 [.02, .35]	.05 [-.16, .25]
Openness	.11 [-.03, .26]	.03 [-.13, .19]
Conscientiousness	-.19 [-.39, .01]	.05 [-.12, .21]
Extraversion with		
Openness	.27 [.15, .39]	.31 [.17, .44]
Conscientiousness	.19 [.05, .32]	.22 [.08, .37]
Openness with		
Conscientiousness	-.15 [-.29, -.02]	.04 [-.10, .18]

Note. CG = control group; EG = experimental group;

covariance structure due to the experimental manipulation.

Effects of socially desirable responding on the mean structure were also investigated. As shown in Table 5 modelling socially desirable responding could revert pre-existing mean differences for most personality factors. Mean differences between groups disappeared for Emotional Stability, Extraversion, and Conscientiousness. Openness, however, could not be modelled accurately by the common method variance factors, as the mean difference still remained with at least a small effect size, $d = -0.35 [-0.56, -0.13]$. The negative sign of the mean difference indicates that the common mean difference for the second measurement occasion (CMV2) overestimated the effect of socially desirable responding for Openness. As expected, the common mean difference was large, $d = 1.18 [1.01, 1.34]$.

Correlations between common method variance factors within groups revealed high linear association of the factors in the control group, $r = .83 [.74, .92]$. The estimated correlation coefficient in the experimental group was close to zero, $r = -.08 [-.22, .06]$.

To assess convergent validity predicted factor scores for each participant for both common method variance factors were correlated with their respective SDS-17 and BIDR test scores. The correlations were very similar for both groups at the first measurement occasion for the SDS-17 (CG: $r = .30 [.19, .49]$, EG: $r = .29 [.18, .39]$), and the BIDR Self-Deceptive Enhancement scale (CG: $r = .52 [.44, .60]$, EG: $r = .44 [.34, .53]$). Expectedly, the BIDR Impression Management scale did not show strong association with the common method variance factor scores (T1), as the scale is constructed to measure deliberate faking of the questionnaire (CG: $r = .13 [.02, .23]$, EG: $r = .15 [.04, .26]$). For the control group this still held true at the second measurement occasion, $r = .11 [.00, .23]$. An at least moderately strong correlation was found between the common method variance factor score (T2) and

Table 5. Tests of mean differences in the SEM.

Factor	CG	EG	<i>p</i>	<i>d</i>	95-% CI
Emotional Stability	0	-0.04	.517	-0.08	[-0.32, 0.16]
Extraversion	0	0.11	.144	0.17	[-0.06, 0.38]
Openness	0	-0.21	.001	-0.35	[-0.56, -0.13]
Conscientiousness	0	0.05	.416	0.09	[-0.12, 0.30]
Common method variance (T1)	0.88	0.88	.165	–	–
Common method variance (T2)	0	1.18	< .001	1.18	[1.01, 1.34]

Note. CG = control group; EG = experimental group; *d* = standardized mean difference; The *p*-value corresponds to the test $H_0 : \mu = 0$ against a two-sided alternative. For all but CMV (T1) this is equivalent to a test of mean differences.

the Impression Management scale in the experimental group, $r = .40$ [.30, .49], further supporting that socially desirable responding is measured by the common method variance factors. A full list of correlations (including correlations across measurement occasions) is given in Table 9 (Appendix B).

3.3 Finite mixture modelling of socially desirable responding

3.3.1 Model selection and goodness-of-fit

Multiple models with increasing number of latent classes were fit to the data of the first measurement occasion to investigate different response patterns in the NEO-FFI based on responses in the SDS-17⁴. Model selection based on information criteria (BIC, aBIC) and BLRT *p*-value both agreed to retain the model with 2 latent classes (see Table 6). Calculation of Bayes Factor also showed that a 2-class model is the more likely model compared to the model with 3 latent classes, $\widehat{BF}_{2,3} > 10$.

⁴ In a first analysis both BIDR scales were also included as predictors of the latent class factor in the models. The results were qualitatively similar to the final analysis but showed non-significant coefficients for both BIDR scales, which were subsequently dropped from the analysis to improve computational stability.

Table 6. Finite mixture model comparison.

# of Classes	AIC	BIC	aBIC	BLRT <i>p</i> -value	Entropy
1	7376.46	7472.52	7402.68	–	–
2	5847.66	6031.05	5897.72	.000	.696
3	5829.03	6108.49	5905.32	.980	.718

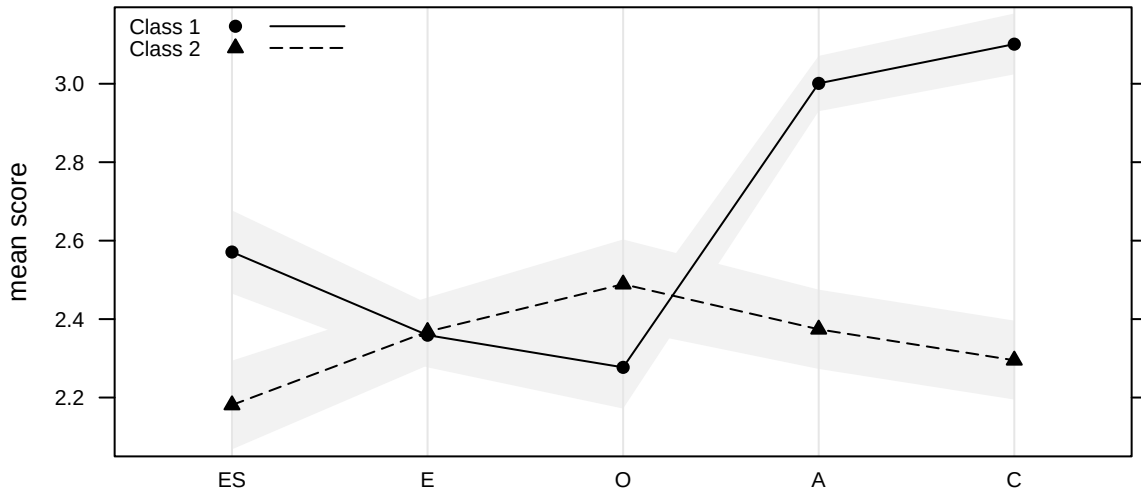


Figure 7. Latent class means and their 95% confidence intervals.

3.3.2 Class comparisons

The two classes were of similar size (Class 1: $n = 311$, Class 2: $n = 271$) and differed in their mean score of Emotional Stability (Class 1: 2.57 [2.47, 2.68], Class 2: 2.18 [2.07, 2.29]), Agreeableness (Class 1: 3.00 [2.93, 3.07], Class 2: 2.38 [2.28, 2.48]), and Conscientiousness (Class 1: 3.10 [3.03, 3.18], Class 2: 2.30 [2.20, 2.38]). No differences in Extraversion (Class 1: 2.36 [2.28, 2.44], Class 2: 2.37 [2.28, 2.46]) or Openness (Class 1: 2.28 [2.17, 2.38], Class 2: 2.49 [2.38, 2.60]) could be observed between latent classes (see also Figure 7 for a graphical representation of latent personality profiles for each class). Social desirability as measured by the SDS-17 mean score was a predictor of latent class membership, $b = 2.99, se = 0.33, p < .001$. With higher SDS-17 mean score the odds of being a member of class 1 were increasingly likely, $OR = 19.83 [10.30, 38.13]$, which shows that higher social desirability scores coincide with higher scores in Emotional Stability, Agreeableness, and Conscientiousness under natural conditions.

Latent classes were also compared with regards to selected proxy variables. Individual participants were assigned a group based on their most likely latent class membership. Group comparisons showed differences in age, $b = -8.01, se = 1.11, p < .001$, and sex, $b = -0.42, se = 0.17, p = .012$. Participants in class 1 were, on average, older ($M = 38.41, SD = 14.20$) than in class 2 ($M = 30.40, SD = 12.34$) and the odds of being female were higher in class 1, $OR = 1.52 [1.10, 2.11]$. Educational level was the same in both classes ($Md = Matura$).

4 Discussion

The present master thesis explored effects of social desirability on personality questionnaires, research context, and individual differences. The results showed that social desirability bias has an effect on the questionnaire and context of a scientific study. A causal effect of general social desirability instructions on both the covariance structure and the mean structure of the NEO-FFI was found. When modelling the data using structural equation models with common method variance factors (Ziegler & Buehner, 2009), it could be shown that trait intercorrelations mostly disappeared. However, substantial correlations between Extraversion and Openness, as well as Extraversion and Conscientiousness remained. Modelling of the mean structure revealed that group differences based on the social desirability instruction could be reverted and represented as a common mean difference caused by social desirable responding. Only a small group difference in Openness remained. Further it could be demonstrated that naturally occurring socially desirable responding was able to differentiate between two groups of responders and therefore expands findings of subpopulation-specific effects. High mean scores in both social desirability and personality traits were given predominantly by older and female participants, whereas low mean scores coincided with younger age of the participants and higher probability of being male.

4.1 Effects of questionnaire and context

Application of Ziegler and Buehner's (2009) method of modelling socially desirable responding yielded similar results to the original study. It can be concluded that the originally proposed approach is appropriate to model socially desirable response bias. However, differences in the results of Ziegler and Buehner and the current master thesis should be addressed.

In the original study it was shown that four out of five Big Five personality traits were affected by socially desirable responding (Neuroticism, Extraversion, Agreeableness, and Conscientiousness). The same number of traits was found to be faked in the current study, but did not include Agreeableness (Emotional Stability (Neuroticism), Extraversion, Openness, and Conscientiousness were affected). The amount of response bias across studies was similar for Neuroticism, and Conscientiousness, but differed substantially for the other personality traits. Differences in the magnitude of socially desirable responding occurred

for Extraversion, where large effects were found in the current thesis, but not by Ziegler and Buehner (2009) who observed only a moderate effect. Openness did not seem to be affected in the original study by Ziegler and Buehner, whereas a small effect was found in the present master thesis. Finally, Agreeableness exhibited a small effect only in the study by Ziegler and Buehner. No effect was found in the current master thesis. These differences were expected and are readily explained by a major factor: The experimental manipulation. In contrast to the master thesis at hand, Ziegler and Buehner used an applicant instruction to evoke socially desirable responding (cf. Ziegler & Buehner, 2009, p. 553). Since socially desirable responding is largely dependent on environmental or contextual factors, differences are to be expected for varying questionnaire instructions (see e.g., Holden, 2007). Additionally, the smaller sample size of Ziegler and Buehner's study ($N = 341$) compared to the current master thesis ($N = 582$) could have been a detrimental factor in finding small effects of social desirability.

Effects of the experimental manipulation on the covariance structure of the questionnaire could also be replicated. Under normal circumstances (assessment at the first measurement occasion) personality traits were moderately correlated, therefore confirming a well established result in Big Five personality research (Van der Linden et al., 2010). The general social desirability instruction then led to even greater correlations between personality traits at the second measurement occasion. This result shows that covariance between Big Five personality traits can be caused by socially desirable responding. Due to the fact that socially desirable responding constitutes systematic measurement error in the form of common method variance between questionnaire items (Podsakoff et al., 2003), high correlations between traits also indicate that higher level personality factors might – at least partially – be a consequence of socially desirable responding.

Harder evidence was provided by modelling socially desirable responding directly. When accounting for socially desirable response bias, pre-existing mean differences between groups disappeared for Emotional Stability, Extraversion, and Conscientiousness, leaving only a small residual group difference in Openness. Thus, a majority of the mean differences could be explained by a single common mean difference. Results are again very similar to the findings by Ziegler and Buehner (2009), who reported a small residual mean difference in Conscientiousness when accounting for socially desirable responding.

By modelling social desirability, correlations between traits could also be reduced. Sub-

stantial correlations between Extraversion and other traits remained (control group: Emotional Stability, Openness, Conscientiousness; experimental group: Openness, Conscientiousness). These findings show that social desirability is a major causal factor for Big Five personality trait intercorrelations. Within social desirability research this is not new (e.g., Pauls & Crost, 2005), but has strong implications on relatively recent developments in personality research. Existence of the General Factor of Personality is dependent on the covariance between Big Five personality traits (Musek, 2007). If correlations between traits are due to methodological artefacts such as social desirability, then the General Factor of Personality itself must be a method factor and not a substantial personality trait. The results of the current master thesis partly support the claims of previous research in that regard (Ziegler & Buehner, 2009), showing that most of the personality traits are uncorrelated when accounting for socially desirable responding. It can further be argued that under these conditions (low amount covariance left to explain) a substantial General Factor of Personality seems unlikely to exist. Even if such a factor could be extracted while accounting for socially desirable response bias, the "generality" (i.e. the representation of all Big Five personality traits in a single factor) of such an hierarchical factor is not guaranteed, as it is almost exclusively comprised of the covariance of Extraversion with other traits.

4.2 Person-specific effects

Modelling of non-experimental data (data of the first measurement occasion) demonstrated that socially desirable responding does not occur homogeneously within a population. In line with previous research concerned with latent classification, the current results support the idea that social desirability is not only dependent on questionnaire and context effects, but also a consequence of individual differences (Maij-de Meij et al., 2008). The two resulting classes differed with respect to Emotional Stability, Agreeableness, and Conscientiousness, but not other Big Five personality traits. Post-hoc analyses of the latent classes also revealed differences in age and sex.

Age-related effects in personality research are well understood and usually higher age is associated with lower Neuroticism (higher Emotional Stability), higher Agreeableness, and higher Conscientiousness (McCrae, Martin, & Costa, 2005; Roberts, Walton, & Viechtbauer, 2006). These personality changes were originally assumed to be caused by better emotional regulation and maturity and only recently investigated and discussed in the context of so-

cial desirability (Soubelet & Salthouse, 2011; Vigil-Colet, Morales-Vives, & Lorenzo-Seva, 2013). High Emotional Stability, Agreeableness, and Conscientiousness were also related with higher age in the present sample, therefore confirming the effect of age on personality measures and strengthening the evidence in favour of social desirability of a possible contribution to the effect (Ones & Viswesvaran, 1998; but see Section 4.3).

Sex differences in social desirability were also observed in the latent classes and showed that female participants scored higher on Emotional Stability, Agreeableness, and Conscientiousness compared to male participants. Cross-cultural research in personality also found sex differences in the Big Five personality traits and the effects were most pronounced for Neuroticism (Emotional Stability), Agreeableness, and Conscientiousness – especially in western cultures (North America, Europe; Schmitt, Realo, Voracek, & Allik, 2008). Although Schmitt et al. (2008) argued that social desirability is an unlikely explanation for sex differences in self-reported personality, the current master thesis supports the social desirability hypothesis (but see Section 4.3). However, the results have to be interpreted cautiously as meta-analytic research in social desirability suggests a reversed effect of sex in applicant settings (i.e. men usually score higher on social desirability measures; Ones & Viswesvaran, 1998).

4.3 Limitations

The main limitation of the master thesis is the reliance on social desirability scales. Although the experimental procedure by Ziegler and Buehner (2009) originally did not rely on such measurement instruments, no other criterion to assess construct validity of the common method variance factors was available in the current study. The second approach also used social desirability scales (SDS-17) to predict latent class membership and was therefore susceptible to errors. The use of social desirability scales is problematic, because they usually have weak psychometric properties. Reliability measures are generally low (Li & Bagger, 2007). Tran et al. (2012) have also shown that one-dimensionality does not hold and questioned if standard factor analytic procedures are even applicable to such a heterogeneous construct. However, the main issue with social desirability scales arises from their lack of divergent validity. They are substantially correlated with Big Five personality measures (Musch et al., 2002; Stöber, 2001), which complicates interpretation of social desirability effects on personality assessments. In the finite mixture modelling approach the

problems of validity and reliability are particularly relevant, because the latent classification was dependent on social desirability scale scores. It can not be ruled out that the resulting classification is caused by the measurement of personality traits by social desirability scales and thus would be a methodological artefact. For the modelling of experimental data the use of social desirability scales as an external criterion to assess convergent validity is not that problematic, but other approaches are certainly preferred (e.g., using the model to predict a distal outcome; see Ziegler & Buehner, 2009).

Further, both the study by Ziegler and Buehner (2009) and the current master thesis showed similar misfit related to the mean structure of the structural equation model. In both studies residual group differences remained, which could not be explained by the mean difference of the common method variance factors. Since the experimental manipulation was the only systematic difference between groups, differences should be explained in its entirety when accounting for socially desirable responding. Hence, the source of this model misfit is unknown. Additional replications are needed to confirm if this type of misfit appears consistently and whether or not the approach needs to be adapted to account for the residual mean differences.

Finally it has to be noted that causal interpretation of the finite mixture modelling findings are not admissible. Neither from the approach used in the current thesis nor related approaches (e.g., Leite & Cooper, 2010) can be concluded with certainty that social desirability is a causal factor for the occurrence of latent classes.

4.4 Conclusion

The present master thesis largely confirmed previous empirical findings and theoretical considerations in social desirability research. It has been shown that questionnaire and context effects, as well as individual differences in socially desirable responding exist and that all three factors should be considered in psychometric and applied research. However, care must be taken when accounting for socially desirable responding in real-world applications (e.g., personnel selection), since use of social desirability scales is often required and experimental procedures are not feasible. It is advised to not rely on social desirability measures, but instead expand experimental methods and to incorporate analysis on individual and subpopulation level in future research.

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, 123–139.
- Asparouhov, T., & Muthén, B. (2006). *Robust chi square difference testing with mean and variance adjusted test statistics*. (Mplus Web Notes: No. 10. May 26, 2006)
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397–438. doi: 10.1080/10705510903008204
- Asparouhov, T., & Muthén, B. (2012). *Using Mplus TECH11 and TECH14 to test the number of latent classes*. (Mplus Web Notes: No. 14. May 22, 2012)
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, 14(4), 317–335.
- Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae*. Göttingen: Hogrefe.
- Burns, G. N., & Christiansen, N. D. (2011). Methods of measuring faking behavior. *Human Performance*, 24(4), 358–372. doi: 10.1080/08959285.2011.597473
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319–333.
- Cervellione, K. L., Lee, Y.-S., & Bonanno, G. A. (2009). Rasch modeling of the Self-Deception Scale of the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement*, 69(3), 438–458. doi: 10.1177/0013164408322020
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation.
- Costa, P. T., & McCrae, R. R. (1989). *NEO PI/FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.

- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349–354.
- Davies, S. E., Connelly, B. S., Ones, D. S., & Birkland, A. S. (2015). The General Factor of Personality: The “Big One,” a self-evaluative trait, or a methodological gnat that won’t go away? *Personality and Individual Differences*, 81, 13–22. doi: <http://dx.doi.org/10.1016/j.paid.2015.01.006>
- DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences*, 33(4), 533–552.
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73(6), 1246–1256.
- Dunkel, C. S., van der Linden, D., Brown, N. A., & Mathes, E. W. (2016). Self-report based General Factor of Personality as socially-desirable responding, positive self-evaluation, and social-effectiveness. *Personality and Individual Differences*, 92, 143–147. doi: <http://dx.doi.org/10.1016/j.paid.2015.12.034>
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Ferrando, P. J. (2005). Factor analytic procedures for assessing social desirability in binary items. *Multivariate Behavioral Research*, 40(3), 331–349. doi: 10.1207/s15327906mbr4003_3
- Goffin, R. D., & Christiansen, N. D. (2003). Correcting personality tests for faking: A review of popular personality tests and an initial survey of researchers. *International Journal of Selection and Assessment*, 11(4), 340–344. doi: 10.1111/j.0965-075X.2003.00256.x
- Higgins, D. M., Peterson, J. B., Pihl, R. O., & Lee, A. G. (2007). Prefrontal cognitive ability, intelligence, Big Five personality, and the prediction of advanced academic and workplace performance. *Journal of Personality and Social Psychology*, 93(2), 298–319.
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92(5), 1270–1285.
- Holden, R. R. (2007). Socially desirable responding does moderate personality scale validity both in experimental and in nonexperimental contexts. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 39(3), 184–201.

- Holden, R. R., & Book, A. S. (2009). Using hybrid rasch-latent class modeling to improve the detection of fakers on a personality inventory. *Personality and Individual Differences*, 47(3), 185–190. doi: <http://dx.doi.org/10.1016/j.paid.2009.02.024>
- Hossiep, R. (2016). Soziale Erwünschtheit. In M. A. Wirtz (Ed.), *Dorsch – Lexikon der Psychologie*. Retrieved from <https://portal.hogrefe.com/dorsch/soziale-erwuenschtheit/>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der „Positive and Negative Affect Schedule“ (PANAS). *Diagnostica*, 42, 139–156.
- Leite, W. L., & Cooper, L. A. (2010). Detecting social desirability bias using factor mixture models. *Multivariate Behavioral Research*, 45(2), 271–293.
- Li, A., & Bagger, J. (2007). The Balanced Inventory of Desirable Responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement*, 67(3), 525–544.
- Lorenzo-Seva, U., & Ferrando, P. J. (2012). A procedure for isolating social desirability variance in a three-way component analysis. *British Journal of Mathematical and Statistical Psychology*, 65(1), 74–88.
- Ludeke, S. G., Weisberg, Y. J., & Deyoung, C. G. (2013). Idiographically desirable responding: Individual differences in perceived trait desirability predict overclaiming. *European Journal of Personality*, 27(6), 580–592.
- Maij-de Meij, A. M., Kelderman, H., & van der Flier, H. (2008). Fitting a mixture item response theory model to personality questionnaire data: Characterizing latent classes and investigating possibilities for improving prediction. *Applied Psychological Measurement*, 32(8), 611–631. doi: 10.1177/0146621607312613
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81–90.
- McCrae, R. R., Martin, T. A., & Costa, P. T. (2005). Age trends and age norms for the NEO Personality Inventory-3 in adolescents and adults. *Assessment*, 12(4), 363–373. doi:

- Musch, J., Brockhaus, R., & Bröder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit. *Diagnostica*, 48, 121–129.
- Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality*, 41(6), 1213–1233.
- Muthén, L. K., & Muthén, B. O. (1998-2010). Mplus user's guide (6th ed.) [Computer software manual]. Los Angeles, CA: Muthén & Muthén.
- Ng, T. W., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58(2), 367–408.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535–569.
- Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11(2–3), 245–269. doi: 10.1080/08959285.1998.9668033
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81(6), 660–679.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598.
- Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR version 6*. Unpublished manuscript, University of British Columbia, Vancouver, Canada.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, J. D. N., & W. D. E. (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Paulhus, D. L., Bruce, M. N., & Trapnell, P. D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality and Social Psychology Bulletin*, 21, 100–108.
- Pauls, C. A., & Crost, N. W. (2005). Effects of different instructional sets on the construct validity of the NEO-PI-R. *Personality and Individual Differences*, 39(2), 297–308. doi: <http://dx.doi.org/10.1016/j.paid.2005.01.003>

- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63(1), 539–569. doi: 10.1146/annurev-psych-120710-100452
- R Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), 1–25.
- Rushton, J. P., & Irwing, P. (2008). A General Factor of Personality (GFP) from two meta-analyses of the Big Five: Digman (1997) and Mount, Barrick, Scullen, and Rounds (2005). *Personality and Individual Differences*, 45(7), 679–683. doi: <http://dx.doi.org/10.1016/j.paid.2008.07.015>
- Saar, K., Aavik, T., & Konstabel, K. (2012). Using principal component scores reduces the effect of socially desirable responding. *Personality and Individual Differences*, 53(3), 279–283.
- Sackheim, H. A., & Gur, R. C. (1978). Self-deception, other-deception and consciousness. In G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (Vol. 2, pp. 139–197). New York: Plenum Press.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94(1), 168–182. doi: <http://dx.doi.org/10.1037/0022-3514.94.1.168>
- Sijtsma, K. (2008). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107–120. doi: 10.1007/s11336-008-9101-0
- Soubelet, A., & Salthouse, T. A. (2011). Influence of social desirability on age differences in self-reports of mood and personality. *Journal of Personality*, 79(4), 741–762. doi: 10.1111/j.1467-6494.2011.00700.x
- Statistik Austria. (2015). *Bildung in Zahlen 2013/14*. Retrieved from <https://www.bmbf>

.gv.at/ministerium/vp/2015/20150422a.pdf?4ut4xd

- Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22, 216–227.
- Stöber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität [The Social Desirability Scale-17 (SDS-17): Development and first findings on reliability and validity]. *Diagnostica*, 45, 173–177.
- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, 17, 222–232.
- Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment*, 78(2), 370–389.
- Tran, U. S., Stieger, S., & Voracek, M. (2012). Psychometric analysis of Stöber's Social Desirability Scale (SDS-17): An item response theory perspective. *Psychological Reports: Measures & Statistics*, 111(3), 870–884.
- Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44(3), 315–327.
- Vickers, A. J., & Altman, D. G. (2001). Analysing controlled trials with baseline and follow up measurements. *BMJ*, 323(7321), 1123–1124. doi: 10.1136/bmj.323.7321.1123
- Vigil-Colet, A., Morales-Vives, F., & Lorenzo-Seva, U. (2013). How social desirability and acquiescence affect the age-personality relationship. *Psicothema*, 25(3), 342–348. doi: 10.7334/psicothema2012.297
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59(2), 197–210.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070.
- Ziegler, M., & Buehner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement*, 69, 548–565. doi: 10.1177/

0013164408324469

Ziegler, M., Maaß, U., Griffith, R., & Gammon, A. (2015). What is the nature of faking? Modeling distinct response patterns and quantitative differences in faking at the same time. *Organizational Research Methods*, 18(4), 679–703.

List of Figures

1	Age distribution in the sample (by group).	18
2	Educational background in the sample (by group).	19
3	Full structural equation model controlling for social desirable responding. . .	23
4	Detail of the SEM illustrating a single personality trait.	25
5	Finite mixture model diagram.	26
6	Boxplot of difference in correlation coefficients across measurement occasions.	30
7	Latent class means and their 95% confidence intervals.	33

List of Tables

1	Reliability measures of the full and short form NEO-FFI scales.	27
2	Means and standard deviations for all variables.	28
3	NEO-FFI correlations with 95% confidence intervals (raw data).	29
4	NEO-FFI correlations with 95% confidence intervals (SEM).	31
5	Tests of mean differences in the SEM.	32
6	Finite mixture model comparison.	32
7	Standardized ESEM factor loadings.	52
8	Standardized SEM factor loadings.	53
9	Correlations between CMV factor scores and measures of social desirability. .	54

Appendix

A Derivation of an item selection criterion

Consider the solution of an exploratory factor analysis or exploratory structural equation model with a set $k = \{1, \dots, m\}$ factors and a set $i = \{1, \dots, n\}$ items. Based on theoretical scale construction the set i is divided in subsets of size $n_1, \dots, n_m$. Within the subsets an item i is assumed to load on a fixed designated factor $j \in k$ with standardized factor loading $\lambda_{ij} \in [-1, 1]$ and cross loads on all other factors $k \setminus j$ with $\lambda_{ik} \in [-1, 1]$.

Item performance is assumed to be high when λ_{ij}^2 is large (close to 1) and $\sum_{k \neq j} \lambda_{ik}^2$ is low (close to 0). Therefore the Item Selection Criterion of item i on a fixed factor j is defined as

$$ISC_{ij} = \frac{\lambda_{ij}^2}{\frac{1}{m-1} \sum_{k \neq j} \lambda_{ik}^2}, \quad \sum_{k \neq j} \lambda_{ik}^2 \neq 0.$$

$ISC_{ij} \in [0, \infty)$, with high values indicating better item performance.

Example

Consider Item 5 and 10 of the NEO-FFI in the current sample (see Table 7). Based on the manual (Costa & McCrae, 1989) the designated factor of both items is Neuroticism ($j = 1$). ESEM with $m = 5$ factors revealed cross loadings for both items (all $\lambda_{ik} \neq 0$). The ISC for both items are

$$ISC_{5,1} = \frac{\lambda_{5,1}^2}{\frac{1}{5-1} \sum_{k \neq 1} \lambda_{5,k}^2} = 49.6$$

and

$$ISC_{10,1} = \frac{\lambda_{10,1}^2}{\frac{1}{5-1} \sum_{k \neq 1} \lambda_{10,k}^2} = 26.28$$

respectively. $ISC_{5,1} > ISC_{10,1}$, and therefore item 5 performs better than item 10 on the basis of their factor loadings.

B Supplementary tables

Table 7. Standardized ESEM factor loadings.

Item	Conscientiousness	Neuroticism	Agreeableness	Extraversion	Openness
5	.62	-.01	.14	-.07	-.08
10	.62	-.20	.12	-.05	-.04
15	.45	-.14	.04	-.09	.04
20	.68	.10	.25	.05	-.04
25	.66	-.08	-.13	.16	-.01
30	.61	-.23	.02	-.07	-.09
35	.65	.07	-.13	.30	.11
40	.61	.00	.22	.01	.03
45	.60	-.02	.20	-.06	-.13
50	.75	-.01	.04	.12	-.09
55	.51	-.37	.16	-.08	.02
60	.59	.08	-.22	.07	.05
1	.06	.46	.03	.05	-.08
6	-.04	.53	.17	-.10	-.13
11	.10	.66	.05	.05	.05
16	-.03	.58	-.08	-.14	.09
21	.10	.70	-.06	-.03	.08
26	-.12	.64	-.01	-.07	.03
31	-.01	.59	.15	-.10	.02
36	.09	.46	-.22	.10	-.09
41	-.26	.63	.03	-.08	-.09
46	.03	.70	-.03	-.18	.14
51	-.21	.70	.04	.07	-.08
56	-.07	.42	.01	.14	.01
4	.12	-.06	.56	.05	.17
9	.06	-.31	.39	-.09	.07
14	.12	.01	.63	-.07	-.03
19	-.04	.22	.40	.02	.06
24	.00	-.25	.53	.05	.10
29	-.12	-.20	.27	.01	.25
34	-.02	-.09	.30	.24	.08
39	.02	.00	.68	.12	-.04
44	-.09	.06	.43	.05	.08
49	.15	.08	.57	-.02	.26
54	.03	-.04	.37	-.20	.14
59	.15	.02	.48	-.15	-.09
2	-.14	.09	.22	.65	-.03
7	.00	.04	.35	.46	.04
12	-.02	-.21	.41	.43	.04
17	.10	.01	.26	.53	.11
22	-.08	-.04	-.16	.60	.07
27	-.14	.00	.31	.26	-.16
32	.13	-.25	-.08	.42	.10

Item	Conscientiousness	Neuroticism	Agreeableness	Extraversion	Openness
37	-.04	-.29	.45	.63	-.06
42	.08	-.44	.33	.37	-.01
47	.12	.26	-.24	.27	.13
52	.24	-.13	-.04	.57	.04
57	.05	-.15	.04	.35	-.06
3	-.27	.14	.11	.01	.32
8	-.11	.03	-.01	-.07	.74
13	.04	.15	.10	.12	.55
18	-.06	-.13	.04	-.17	.40
23	-.03	.10	.04	-.04	.66
28	-.08	-.13	-.02	.07	.29
33	.07	.15	.26	.09	.47
38	-.18	-.08	-.03	-.14	.23
43	-.02	.20	.06	-.02	.65
48	-.05	-.11	.03	.02	.61
53	.18	-.23	-.17	.19	.54
58	-.11	-.08	-.22	.01	.66

Note. Significant loadings are highlighted in bold typeface.

Table 8. Standardized SEM factor loadings.

T	Item	ES		E		O		C		CMV1		CMV2	
		CG	EG	CG	EG	CG	EG	CG	EG	CG	EG	CG	EG
1	1	.47	.45	-	-	-	-	-	-	.16	.17	-	-
	2	.50	.45	-	-	-	-	-	-	.40	.38	-	-
	3	.63	.57	-	-	-	-	-	-	.43	.41	-	-
	4	.23	.23	-	-	-	-	-	-	.61	.65	-	-
	5	.48	.45	-	-	-	-	-	-	.42	.42	-	-
	6	.48	.45	-	-	-	-	-	-	.62	.61	-	-
2	1	.43	.25	-	-	-	-	-	-	-	-	.26	.21
	2	.44	.32	-	-	-	-	-	-	-	-	.49	.67
	3	.61	.38	-	-	-	-	-	-	-	-	.54	.73
	4	.42	.30	-	-	-	-	-	-	-	-	.71	.74
	5	.49	.30	-	-	-	-	-	-	-	-	.37	.45
	6	.43	.26	-	-	-	-	-	-	-	-	.59	.49
1	7	-	-	.72	.64	-	-	-	-	-.02	-.02	-	-
	8	-	-	.71	.61	-	-	-	-	.10	.09	-	-
	9	-	-	.56	.51	-	-	-	-	.11	.11	-	-
	10	-	-	.38	.33	-	-	-	-	.44	.43	-	-
	11	-	-	.53	.45	-	-	-	-	.45	.43	-	-
	12	-	-	.31	.25	-	-	-	-	.31	.28	-	-
2	7	-	-	.75	.43	-	-	-	-	-	-	.02	.50
	8	-	-	.76	.49	-	-	-	-	-	-	.10	.62
	9	-	-	.55	.31	-	-	-	-	-	-	.12	.39
	10	-	-	.52	.28	-	-	-	-	-	-	.38	.72
	11	-	-	.51	.34	-	-	-	-	-	-	.48	.70

T	Item	ES		E		O		C		CMV1		CMV2	
		CG	EG	CG	EG	CG	EG	CG	EG	CG	EG	CG	EG
1	12	-	-	.26	.16	-	-	-	-	-	-	.28	.31
	13	-	-	-	-	.73	.68	-	-	-.20	-.23	-	-
	14	-	-	-	-	.59	.56	-	-	-.08	-.09	-	-
	15	-	-	-	-	.63	.61	-	-	-.18	-.21	-	-
	16	-	-	-	-	.67	.57	-	-	-.27	-.28	-	-
	17	-	-	-	-	.58	.59	-	-	.01	.01	-	-
2	18	-	-	-	-	.65	.59	-	-	-.15	-.15	-	-
	13	-	-	-	-	.74	.53	-	-	-	-	-.12	.35
	14	-	-	-	-	.71	.54	-	-	-	-	-.11	.48
	15	-	-	-	-	.70	.51	-	-	-	-	-.19	.35
	16	-	-	-	-	.71	.55	-	-	-	-	-.23	.32
	17	-	-	-	-	.68	.55	-	-	-	-	-.09	.32
	18	-	-	-	-	.68	.46	-	-	-	-	-.12	.41
1	25	-	-	-	-	-	-	.63	.60	.20	.20	-	-
	26	-	-	-	-	-	-	.58	.50	.42	.39	-	-
	27	-	-	-	-	-	-	.79	.70	.14	.13	-	-
	28	-	-	-	-	-	-	.48	.46	.39	.40	-	-
	29	-	-	-	-	-	-	.68	.56	.25	.22	-	-
	30	-	-	-	-	-	-	.69	.76	.32	.38	-	-
2	25	-	-	-	-	-	-	.59	.50	-	-	.32	.44
	26	-	-	-	-	-	-	.57	.39	-	-	.47	.70
	27	-	-	-	-	-	-	.64	.59	-	-	.35	.57
	28	-	-	-	-	-	-	.45	.29	-	-	.50	.70
	29	-	-	-	-	-	-	.59	.34	-	-	.37	.63
	30	-	-	-	-	-	-	.70	.52	-	-	.42	.54

Note. ES = Emotional Stability, E = Extraversion, O = Openness, CMV1 = Common method variance (T1), CMV2 = Common method variance (T2); Non-significant loadings are highlighted in bold typeface.

Table 9. Correlations between CMV factor scores and measures of social desirability.

	T1		T2	
	CG	EG	CG	EG
CMV (T1) with				
SDS-17	.30 [.19, .40]	.29 [.18, .39]	.27 [.16, .37]	.06 [-.06, .17]
BIDR SDE	.52 [.44, .60]	.44 [.34, .53]	.50 [.41, .58]	-.05 [-.17, .06]
BIDR IM	.13 [.02, .24]	.15 [.04, .26]	.09 [-.03, .20]	.09 [-.03, .20]
CMV (T2) with				
SDS-17	.26 [.15, .37]	.04 [-.07, .16]	.31 [.20, .41]	.51 [.42, .59]
BIDR SDE	.48 [.39, .56]	-.03 [-.14, .09]	.54 [.45, .61]	.52 [.43, .60]
BIDR IM	.12 [.01, .23]	-.09 [-.20, .03]	.11 [.00, .23]	.40 [.30, .49]

Social Desirable Responding

Die Untersuchung soll als **randomisiertes und kontrolliertes Experiment** durchgeführt werden. Untersuchungsteilnehmer/innen sind individuell zu testen und **erst nach der zweiten Testung über den eigentlichen Zweck und Ablauf der Untersuchung zu informieren** (= 2 unterschiedliche Instruktionen zu T2, randomisierte Zuteilung). Jede/r Teilnehmer/in füllt über 2 Zeitpunkte (T1 und T2) insgesamt 2 Fragebögen aus.

Zu T1 ist von **allen** Teilnehmer/innen

Fragebogen-SDR-1

auszufüllen (siehe Abbildung 1). Zwischen T1 und T2 soll ein Abstand von **maximal 7 Tagen** liegen, **optimal** wären **2-5 Tage**. Für jede Versuchsperson ist das **genaue Datum** der Testung am Fragebogen zu vermerken.

Zu T2 beantwortet eine Hälfte der Teilnehmer/innen den Fragebogen noch einmal (diesmal ohne soziodemographische Daten am Schluss; **Fragebogen-SDR-2**). Die andere Hälfte wird instruiert (**Fragebogen-SDR-2a**), den Fragebogen so auszufüllen, wie das ihrer Meinung nach eine von anderen bewunderte und gelobte Person machen würde. Der Fragebogen-SDR-2a enthält noch einmal Fragen zur Soziodemographie. **Diese Fragen sollen ebenfalls unter dem Blickwinkel einer von anderen bewunderten und gelobten Person ausgefüllt werden.** Abbildung 1 verdeutlicht das Design.

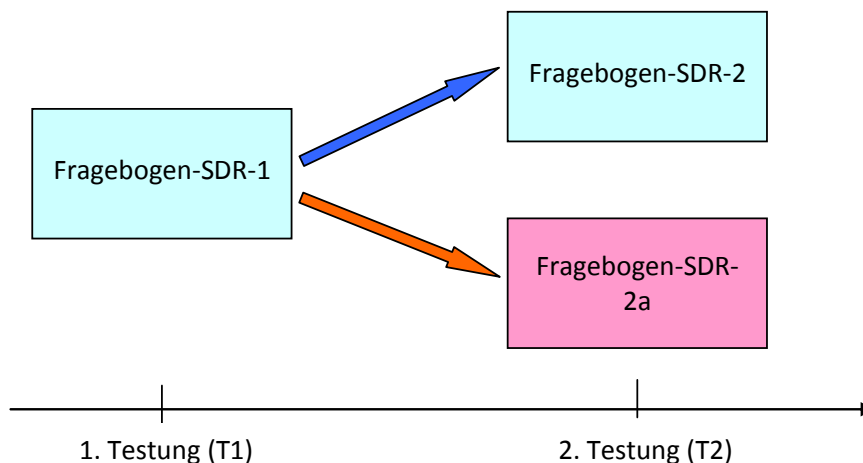


Abbildung 1. Design der Untersuchung.

Die Zuteilung zu den Gruppen mit ehrlicher (Fragebogen-SDR-2) und sozial erwünschter (Fragebogen-SDR-2a) zweiter Testung erfolgt durch eine **Block-Randomisierung**. Am einfachsten ist diese **bereits vor der ersten Testung** mithilfe von Losen durchzuführen.

Die Untersuchung soll im Einzelnen wie folgt durchgeführt werden:

1. Sie fertigen n (= Anzahl aller Personen, die von Ihnen insgesamt untersucht werden) Lose an;
2. $n/2$ dieser Lose legen den Untersuchungsablauf Fragebogen-SDR-1–Fragebogen-SDR-2 (**Untersuchungsablauf 1**) fest;
3. die übrigen $n/2$ Lose legen den Untersuchungsablauf Fragebogen-SDR-1–Fragebogen-SDR-2a (**Untersuchungsablauf 2**) fest;
4. Sie legen sich jeweils $n/2$ Paare von Fragebogen ($n/2$ Paare Fragebogen-SDR-1 und Fragebogen-SDR-2; $n/2$ Paare Fragebogen-SDR-1 und Fragebogen-SDR-2a) bereit;
5. alle Paare von Fragebögen werden von Ihnen nummeriert; jedes Paar erhält eine **andere** Nummer, innerhalb jedes Paares haben beide Fragebögen **dieselbe** Nummer; das gewährleistet die eindeutige Zuordnung der Fragebögen von T1 und T2;
6. **vor der ersten Testung** ziehen Sie für jede/n Teilnehmer/in **unbeobachtet** ein Los aus ihrem Lospool; das Los entscheidet, ob der/die Teilnehmer/in den Untersuchungsablauf 1 oder 2 durchläuft; **Sie teilen den Teilnehmer/innen weder die Entscheidung per Los noch die Entscheidung selbst mit**;
7. das gezogene Los kommt **nicht** in den Pool zurück, sondern wird im Anschluss **vernichtet**;
8. Teilnehmer/innen füllen Fragebogen-SDR-1 aus und sollen sich – für die zweite Testung – **die Kennzahl am Fragebogen selbst notieren**; Teilnehmer/innen sollen Ihnen bei der zweiten Testung **selbst** sagen, welche Nummer sie hatten;
9. Sie machen sich ggf. selbst eine Notiz, welche Nummer Ihr/e Teilnehmer/in hatte; bei der zweiten Testung wird diese Notiz **vernichtet**;
10. Sie machen sich einen zweiten Testtermin mit dem/r Teilnehmer/in aus (optimal: nach 2-5 Tagen; maximal nach 7 Tagen);
11. Sie legen dem/r Teilnehmer/in zur zweiten Testung den passenden zweiten Fragebogen vor;
12. nach dem Ausfüllen des zweiten Fragebogens klären Sie die Teilnehmer/innen über das eigentliche Design der Studie (2 unterschiedliche Instruktionen zum 2. Testzeitpunkt, randomisierte Zuweisung zu den Untersuchungsbedingungen) auf.

Die Block-Randomisierung gewährleistet 1) dass Personen **zufällig** zu den beiden Bedingungen (Untersuchungsablauf 1 und 2) zugewiesen werden und 2) dass die resultierenden Gruppen **gleichgroß** sind. Die Erhebung zu T1 erlaubt die Messung einer **Baseline**, die Erhebung zu T2 die spezifische Untersuchung des Effekts der Instruktion zu T2.

SE Bachelorarbeit I

Social Desirable Responding

Überblick über die verwendeten Verfahren

Der Fragebogen enthält folgende Instrumente in der angegebenen Reihenfolge:

1. NEO-Fünf-Faktoren Inventar (NEO-FFI; Costa & McCrae, 1989; dt. Übersetzung: Borkenau & Ostendorf, 1993)
2. Fragebogen zur Geschlechtstypizität (GTS+; Altstötter-Gleich, 2004)
3. Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988; dt. Übersetzung: Krohne, Egloff, Kohlmann, & Tausch, 1996; Instruktion „im Allgemeinen“)
4. Behavioral Inhibition System/Behavioral Approach System Scales (BIS/BAS; Carver & White, 1994; dt. Übersetzung: Strobel, Beauducel, Debener, & Brocke, 2001)
5. Balanced Inventory of Social Desirable Responding (BIDR; Paulhus, 1994; dt. Übersetzung: Musch, Brockhaus, & Bröder, 2002; Teil 5, Items 1-20)
6. Soziale-Erwünschtheits-Skala-17 (SES-17; Stöber, 1999; Antwortformat nicht dichotom, sondern 7-stufig wie im BIDR; Teil 5, Items 21-37)
7. Soziodemografische Daten

Messinstrumente:

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, 123-139.

Borkenau, P., & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) nach Costa und McCrae*. Göttingen: Hogrefe.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319-333.

Costa, P. T., & McCrae, R. R. (1989). *The NEO PI/FFI manual supplement*. Odessa, Florida: Psychological Assessment Resources.

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der „Positive and Negative Affect Schedule“ (PANAS). *Diagnostica*, 42, 139-156.

Musch, J., Brockhaus, R., & Bröder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit. *Diagnostica*, 48, 121-129.

Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR version 6*. Unpublished manuscript, University of British Columbia, Vancouver, Canada.

Stöber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität [The Social Desirability Scale-17 (SDS-17): Development and first findings on reliability and validity]. *Diagnostica*, 45, 173-177.

Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22, 216-227.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.

Stichprobe:

Allgemeinbevölkerung

Gesamt-N: 400+ Probanden (20+/StudentIn)
Geschlecht: 50% Männer, 50% Frauen
Alter: 18 bis 65 Jahre
Bildung: alle Bildungsschichten (aber: Personen mit akademischer Bildung < 25%; vgl. Statistik Austria, 2011: nur 14.1% der österreichischen Bevölkerung im Alter von 25 bis 64 Jahren haben einen universitären oder tertiären Abschluss)

Deutsch als Muttersprache ist keine explizite Voraussetzung für die Teilnahme; Deutsch muss aber **ausreichend beherrscht** werden, damit sichergestellt ist, dass die Fragebögen richtig verstanden und ausgefüllt werden.

D Survey instructions for participants

Nummer: _____ Datum: _____

Universität Wien
Fakultät für Psychologie

Information zur Studienteilnahme Fragebogen 1

Wir studieren Psychologie an der Universität Wien und laden Sie herzlich dazu ein, an einer wissenschaftlichen Studie teilzunehmen. Die Studie befasst sich mit der Stabilität von Persönlichkeitsmerkmalen. Aus diesem Grund wird Ihnen neben diesem Fragebogen in einigen Tagen noch ein weiterer, ähnlicher Fragebogen vorgelegt werden.

Der folgende Fragebogen beinhaltet verschiedene Fragen zu Ihren persönlichen Einschätzungen bzw. Empfindungen. Bitte lesen Sie sich die Instruktionen der Fragebögen genau durch und beantworten Sie alle Fragen, selbst wenn manche ähnlich klingen.

Ihre Rechte:

Selbstverständlich können Sie jederzeit vor und während der Studie weitere Informationen über Zweck, Ablauf usw. der Studie von den Personen erfragen, die die Datenerhebung durchführen. Die Teilnahme an dieser Studie ist vollkommen freiwillig. Sie können die Untersuchung jederzeit auch ohne Angabe von Gründen abbrechen.

Datenschutz:

Sämtliche in dieser Studie erhobenen Daten werden streng vertraulich behandelt und nicht an Dritte weitergegeben. Die Kennzahl am Kopf dieser Seite dient der Zuordnung dieses Fragebogens zum zweiten Fragebogen, der Ihnen in einigen Tagen vorgelegt wird. Ihre Anonymität bleibt jedoch stets gewahrt. Die Daten werden ausschließlich gruppenbezogen analysiert; es findet keine personenbezogene Auswertung statt.

Einverständniserklärung:

Durch das Ankreuzen des unten stehenden Kästchens bestätigen Sie, dass Sie das vorliegende Informationsblatt gelesen und verstanden haben. Sie erklären sich mit der Teilnahme an dieser Studie sowie mit der Analyse Ihrer Daten durch befugte Personen einverstanden.

☐ Ich habe die Information gelesen und möchte an dieser Studie teilnehmen

Vielen Dank für Ihre Teilnahme!

Diese Studie wird im Rahmen einer Lehrveranstaltung der Fakultät für Psychologie der Universität Wien unter Leitung von Mag. Dr. Ulrich Tran im Sommersemester 2013 durchgeführt.

Mag. Dr. Ulrich Tran
Universität Wien, Fakultät für Psychologie
Institut für Psychologische Grundlagenforschung
und Forschungsmethoden
Liebiggasse 5, A-1010 Wien
E-Mail: ulrich.tran@univie.ac.at

Nummer: _____

Datum: _____

Universität Wien
Fakultät für Psychologie

Information zur Studienteilnahme Fragebogen 2

Sie haben bereits vor einigen Tagen einen Fragebogen zu verschiedenen persönlichen Einschätzungen bzw. Empfindungen ausgefüllt. Nun bitten wir Sie, diesen zweiten Fragebogen auszufüllen, der dieselben Fragen noch ein weiteres Mal enthält.

Bitte lesen Sie sich wiederum die Instruktionen der Fragebögen genau durch und beantworten Sie alle Fragen, selbst wenn manche ähnlich klingen.

Ihre Rechte:

Selbstverständlich können Sie jederzeit vor und während der Studie weitere Informationen über Zweck, Ablauf usw. der Studie von den Personen erfragen, die die Datenerhebung durchführen. Die Teilnahme an dieser Studie ist vollkommen freiwillig. Sie können die Untersuchung jederzeit auch ohne Angabe von Gründen abbrechen.

Datenschutz:

Sämtliche in dieser Studie erhobenen Daten werden streng vertraulich behandelt und nicht an Dritte weitergegeben. Die Kennzahl am Kopf dieser Seite dient der Zuordnung dieses Fragebogens zum ersten Fragebogen, der Ihnen bereits vor einigen Tagen vorgelegt wurde. Ihre Anonymität bleibt jedoch stets gewahrt. Die Daten werden ausschließlich gruppenbezogen analysiert; es findet keine personenbezogene Auswertung statt.

Einverständniserklärung:

Durch das Ankreuzen des unten stehenden Kästchens bestätigen Sie, dass Sie das vorliegende Informationsblatt gelesen und verstanden haben. Sie erklären sich mit der Teilnahme an dieser Studie sowie mit der Analyse Ihrer Daten durch befugte Personen einverstanden.

☐ Ich habe die Information gelesen und möchte an dieser Studie teilnehmen

Vielen Dank für Ihre Teilnahme!

Diese Studie wird im Rahmen einer Lehrveranstaltung der Fakultät für Psychologie der Universität Wien unter Leitung von Mag. Dr. Ulrich Tran im Sommersemester 2013 durchgeführt.

Mag. Dr. Ulrich Tran
Universität Wien, Fakultät für Psychologie
Institut für Psychologische Grundlagenforschung
und Forschungsmethoden
Liebiggasse 5, A-1010 Wien
E-Mail: ulrich.tran@univie.ac.at

Nummer: _____

Datum: _____

Universität Wien
Fakultät für Psychologie

Information zur Studienteilnahme Fragebogen 2a

Sie haben bereits vor einigen Tagen einen Fragebogen zu verschiedenen persönlichen Einschätzungen bzw. Empfindungen ausgefüllt. Nun bitten wir Sie, diesen zweiten Fragebogen auszufüllen, der dieselben Fragen noch ein weiteres Mal enthält.

Überlegen Sie sich dieses Mal jedoch, wie eine Person die Fragen beantworten würde, die von anderen bewundert und gelobt wird. Kreuzen Sie bitte alle Antworten in diesem Fragebogen entsprechend dieser Vorstellung an. Uns interessiert, wie sehr Angaben in Fragebögen durch sozial erwünschtes Antworten verzerrt werden können.

Bitte lesen Sie sich wiederum die Instruktionen der Fragebögen genau durch und beantworten Sie alle Fragen, selbst wenn manche ähnlich klingen.

Ihre Rechte:

Selbstverständlich können Sie jederzeit vor und während der Studie weitere Informationen über Zweck, Ablauf usw. der Studie von den Personen erfragen, die die Datenerhebung durchführen. Die Teilnahme an dieser Studie ist vollkommen freiwillig. Sie können die Untersuchung jederzeit auch ohne Angabe von Gründen abbrechen.

Datenschutz:

Sämtliche in dieser Studie erhobenen Daten werden streng vertraulich behandelt und nicht an Dritte weitergegeben. Die Kennzahl am Kopf dieser Seite dient der Zuordnung dieses Fragebogens zum ersten Fragebogen, der Ihnen bereits vor einigen Tagen vorgelegt wurde. Ihre Anonymität bleibt jedoch stets gewahrt. Die Daten werden ausschließlich gruppenbezogen analysiert; es findet keine personenbezogene Auswertung statt.

Einverständniserklärung:

Durch das Ankreuzen des unten stehenden Kästchens bestätigen Sie, dass Sie das vorliegende Informationsblatt gelesen und verstanden haben. Sie erklären sich mit der Teilnahme an dieser Studie sowie mit der Analyse Ihrer Daten durch befugte Personen einverstanden.

☐ Ich habe die Information gelesen und möchte an dieser Studie teilnehmen

Vielen Dank für Ihre Teilnahme!

Diese Studie wird im Rahmen einer Lehrveranstaltung der Fakultät für Psychologie der Universität Wien unter Leitung von Mag. Dr. Ulrich Tran im Sommersemester 2013 durchgeführt.

Mag. Dr. Ulrich Tran
Universität Wien, Fakultät für Psychologie
Institut für Psychologische Grundlagenforschung
und Forschungsmethoden
Liebiggasse 5, A-1010 Wien
E-Mail: ulrich.tran@univie.ac.at

E Curriculum vitae

Philipp GEWESSLER

ANGABEN ZUR PERSON

GEBURTSORT UND -DATUM: Linz, Österreich | 21. Mai, 1991
ADRESSE: Währinger Gürtel 148/11-12, Wien, Österreich
TELEFON: +43 676 70 29 238
E-MAIL: philipp.gewessler@univie.ac.at

BERUFSERFAHRUNG

OKTOBER 2014 - AUGUST 2016	Studentischer Mitarbeiter (Studienassistent) <i>Fakultät für Psychologie, Universität Wien</i> Unterstützung bei der Vorbereitung und Durchführung von Lehrveranstaltungen und Prüfungen, Lehrunterstützung durch Tutorien, sowie Mithilfe bei aktuellen Forschungsarbeiten im Bereich der psychologischen Statistik am Institut für psychologische Grundlagenforschung und Forschungsmethoden: Forschungsmethoden der Psychologie bei Univ.-Prof. Dr. Alexander VON EYE (bis Sommersemester 2015) und Assoz. Prof. MMag. DDDr. Martin VORACEK, Privatdoz. (ab Wintersemester 2016).
AUGUST 2014 - OKTOBER 2014	Psychologisches Pflichtpraktikum <i>Fakultät für Psychologie, Universität Wien</i> Mitarbeit bei einer Monte-Carlo Simulationsstudie zum Thema Effektrichtungen im Kontext von Regression und Strukturgleichungsmodellen (inkl. programmieren der Simulation in R) am Institut für psychologische Grundlagenforschung und Forschungsmethoden: Forschungsmethoden der Psychologie.
MÄRZ 2014 - JULI 2014	Studentischer Mitarbeiter (Studienassistent) <i>Fakultät für Psychologie, Universität Wien</i> Unterstützung bei der Vorbereitung und Durchführung von Lehrveranstaltungen und Prüfungen sowie bei der Betreuung von Studierenden, insbesondere bei der Diplomarbeitsbetreuung, am Institut für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung bei ao. Univ.-Prof. Dr. Germain WEBER.
JULI 2013 - JULI 2014	Wissenschaftlicher Projektmitarbeiter <i>Fakultät für Psychologie, Universität Wien</i> Projekt: "ESTSS", Projektleitung: Ass.-Prof. Dr. Brigitte LUEGER-SCHUSTER. Der Aufgabenbereich umfasste hauptsächlich die Organisation des Zertifizierungsprogrammes der ESTSS (European Society for Traumatic Stress Studies).
OKTOBER 2010 - <i>aktuell</i>	Web Administrator <i>fsms gmbh</i> Programmieren und instand halten der Firmen-Website.
2010, 2011, 2012	Betreuer für Menschen mit Beeinträchtigung <i>Ferienaktion, Volkshilfe lebensART GmbH</i> Projekt-basierte Anstellung in den Sommermonaten.
AUGUST 2009 - APRIL 2010	Zivildienst <i>Volkshilfe Oberösterreich</i> Betreuung von Menschen mit Beeinträchtigung in einer Wohneinrichtung.

KONFERENZBEITRÄGE

2016 Gewessler, P., Voracek, M., & Tran, U. S. (2016). *Ein Mehrmethodenansatz zur Beurteilung des Einflusses von sozial erwünschtem Antwortverhalten auf die Big Five*. 50. Kongress der Deutschen Gesellschaft für Psychologie, 18.–22. September 2016, Universität Leipzig, Deutschland.

PUBLIKATIONEN (IN ARBEIT)

Gewessler, P., & Wiedermann, W. (in prep.). Direction of effects in Actor-Partner Interdependence Models for distinguishable dyads.

Gewessler, P., & Wiedermann, W. (in prep.). Sample-size determination for directional dependence tests in mediation models.

SCHUL- UND BERUFSBILDUNG

OKTOBER 2014 Bachelorstudium Statistik, **Universität Wien**, Wien

OKTOBER 2013 Masterstudium Psychologie, **Universität Wien**, Wien

Schwerpunkt: "Psychologische Grundlagen: Geist & Gehirn".

Die inhaltlichen Schwerpunkte des Studiums liegen in der biologischen Psychologie (neuronale Korrelate von Empathie und prosozialem Verhalten), der Kognitionspsychologie (visuelle Aufmerksamkeit) und der allgemeinen Psychologie (psychologische Ästhetik).

Masterthesis: *Statistical Modeling of Social Desirability Bias in Personality Questionnaires*. [Arbeitstitel]

AUGUST 2013 Bachelor of Science (BSc), **Universität Wien**, Wien

Bachelorarbeit I: Mindfulness | IV-Leiter: Mag. Dr. Ulrich S. TRAN

Bachelorarbeit II: Rapid Automatized Naming | IV-Leiter: ao. Univ.-Prof. Mag. Dr. Alfred SCHABMANN

OKTOBER 2010 Bachelorstudium Psychologie, **Universität Wien**, Wien

JUNI 2009 Matura, **BG/BRG Ramsauerstraße**, Linz

SPRACHEN

DEUTSCH: Muttersprache

ENGLISCH: fließend

SPANISCH: grundlegende Fähigkeiten

COMPUTERKENNTNISSE

GRUNDLEGENDE KENNTNISSE: CSS, HTML, Linux, MATLAB, Javascript (D3.js).

FORTGESCHRITTENE KENNTNISSE: L^AT_EX, Microsoft Office, M₊PLUS, R, SPSS.