Empathy, Narcissism and Gaslighting: Establishing Correlations and Investigating Moderation Effects

Tom Strombach (s2813459)

Department of Behavioural, Management and Social Sciences

M12 BSc Thesis PSY

Dr. Peter M. ten Klooster

Dr. Jannis T. Kraiss

June 19, 2024

Abstract

Background

Gaslighting refers to manipulative behaviours aimed at undermining a victim's perception of reality. This study investigates the relationship between empathy and acceptance of gaslighting, considering narcissism as a potential moderator.

Methods

The Toronto Empathy Questionnaire (TEQ) and the Brief Pathological Narcissism Inventory (B-PNI) assessed empathy and narcissism, respectively. Acceptance of gaslighting was measured using a recently developed questionnaire by March et al. (2023), whose psychometric properties were re-evaluated with a confirmatory factor analysis. A moderation analysis was conducted using the PROCESS macro for RStudio.

Results

Out of 147 participants, 96 provided sufficient data. Empathy had a significant negative relation with acceptance of gaslighting (d = -0.83, p = .001, r = .40). Both vulnerable (d = 0.27, p < .001, r = .43) and grandiose (d = 0.12, p = .003, r = .37) narcissism, as well as generalised narcissism (d = 0.23, p = .001, r = .40), had significant positive relations. Interaction effects between empathy and narcissism on gaslighting acceptance were insignificant. The acceptance of gaslighting questionnaire showed good item measurement quality despite inadequate model fit for strict unidimensionality.

Conclusion

Higher empathy scores correlated with lower acceptance of gaslighting, while higher narcissism scores correlated with higher gaslighting acceptance. No moderating effect of narcissism was found. The gaslighting acceptance scale sufficiently measures the construct despite some psychometric limitations.

Introduction

In his 1938 thriller play “Gas Light” Patrick Hamilton paints the picture of a less than happy marriage, riddled with deception, manipulation and psychological terror. In the play, a husband tries to convince his wife that she has lost her mind in order to get her institutionalised so he can steal precious jewellery hidden in the attic of their family home. To destabilise his wife, the husband repeatedly lies to her and the people close to her in order to undermine her grip on reality. The play was adapted by filmmakers twice, and with its second iteration the term *gaslighting* started to enter the mainstream.

Today this concept is widely accepted as a form of manipulation and investigated by philosophers, psychologists and sociologists. Gaslighting is especially prevalent in romantic relationships, and relationships with unequal power dynamics (Klein et al., 2023; March et al., 2023; Peeren et al., 2024; Stark, 2019). Nonetheless, there is a considerable lack of research on the topic of gaslighting and articles may still contradict each other on matters as fundamental as its definition. However, for the most part the definitions found in the literature mirror the aspects that were already present in the original play by Hamilton. Gaslighting has been described as a form of “wrongful manipulation and [...] emotional abuse” (p. 221) that often includes covert types of manipulation like deception, denial, and contradiction (March et al., 2023; Stark 2019).

These attributes make it hard to investigate important aspects of gaslighting on the whole, its prevalence and predictors. Due to its discrete nature, it is likely underreported by (oftentimes female) victims and offenders, while the consequences can be detrimental (Stark, 2019). Gaslighting usually constitutes continued attacks on the victims’ self-confidence, especially related to their own mental capabilities (March et al., 2023; Sodoma, 2022). Victims of gaslighting might therefore avoid seeking help because they have been mentally destabilised to a point where they do not trust their own judgements or the people around them (March et al., 2023). While there is significant overlap with other types of emotional abuse, March et al. (2023) argues that a characteristic feature of gaslighting is the “doubting of one’s reality” (p. 2). In addition, it can be said that gaslighting mainly occurs in close relationships with unequal power dynamics (March et al., 2023; Stark, 2019), wherein the “target’s trust and emotional investment in the relationship [are used] as leverage” (Sodoma, 2022, p. 322).

Considering the sinister nature of this type of manipulation, it seems necessary to investigate the phenomenon and to find ways to reduce the occurrences of gaslighting. Therefore, it must be mentioned that the measurement of anti-social behaviours in-vivo as well as in surveys is notoriously difficult and highly susceptible to social desirability effects (Ferrer-Perrez, 2020). For this reason, March et al. (2023) chose a different approach in which they tried to investigate attitudes rather than actual behaviours. In their study, they

developed and validated a short questionnaire with ten items meant to measure participants' *acceptance towards behaviours that are associated with gaslighting*.

The same study by March et al. (2023) established a correlation between dark tetrad traits (i.e. narcissism, sadism, psychopathology and Machiavellianism) and the acceptance of gaslighting behaviours. Specifically, it was found that male participants with a high score on vulnerable narcissism also scored higher than the other groups on acceptance towards gaslighting. In line with these findings, Miano et al. (2021) established the correlation of gaslighting with other aversive personality traits, which rhymes with findings from correlates with intimate partner violence (IPV) and coercive control (Crossman et al., 2015). In terms of demographics, Li and Samp (2023) showed that gaslighters of LGBTQ+ individuals are most likely to be the parents, more precisely cisgender men (i.e. fathers) aged above 45 years. Which already hints at the situations in which gaslighting occurs, which are primarily romantic relationships, but also relationships with unequal power dynamics of various natures, such as parents with their children and business relationships (Klein et al., 2023; Kukreja & Pandey, 2023; Li & Samp, 2023; March et al., 2023). Apart from these established characteristics, after years of studying the phenomenon of gaslighting, most studies on the topic present as "idiosyncratic case studies" (Klein et al., 2023). Therefore, this study aims to provide empirical data on less explored dimensions of gaslighting, specifically its antecedents in regards to the acceptance of gaslighting behaviours.

In order to find potential predictors, the aforementioned characteristic feature of gaslighting behaviours, namely *doubting the reality of the victim,* will serve as a starting point for the present study. This feature relates to the acceptance and understanding of people's subjective cognitive and emotional experience of the world, and with that relates closely to the concept of *empathy* (Sodoma, 2022). In line with previous research and for the purpose of this study *empathy* will be defined as the capacity to engage with another person's point of view, and feel what they are feeling (Cho & Lee, 2023; Dor-Zidermann et al., 2021; Israelashvili et al., 2020; Kappelmayer et al., 2023; Karnaze et al., 2022; Sodoma, 2022; Stevens & Taber, 2021). So far, the relation of empathy and the acceptance of gaslighting behaviours has not yet been explored in the literature. However, it has already been established that empathy has positive effects on anti-social behaviours and coercive control, which closely aligns with gaslighting behaviours (Cho & Lee, 2023; Felton, 2024; Israelashvili et al., 2020; Karnaze et al., 2022). Therefore, a potential negative correlation between scores on empathy and scores on the acceptance of gaslighting will be explored.

In addition to that, March et al. (2023) already pointed out that the acceptance of gaslighting behaviours seems to be related to dark tetrad traits, and especially vulnerable narcissism. Their study also gave a speculative explanation for this relation. According to March et al. (2023), vulnerable narcissists are especially susceptible to "rejection",

“experienc[ing] inadequacy”, and “apprais[ing] negative social feedback as devaluing of the self” (Besser & Priel, 2010; Grieve & March, 2021; Hart et al., 2017; March et al., 2021, as cited in March et al., 2023, p. 8). The authors argue that, therefore, individuals high in vulnerable narcissism are in special need of control of the relationship and their partner's behaviour to avoid “ego-threats” (March et al., 2023, p. 8).

Interestingly, when it comes to empathy and manipulating behaviours, there seems to be an interplay between empathic capacities and *narcissism.* Overall, a negative relationship between narcissism and empathy has already been established (Felton, 2024), but as Heym et al. (2021) pointed out, so called *dark empaths* combine a high score on empathy with high scores on the dark trait traits. Expanding on that, other studies showed that narcissism seems to be related with emotion recognition and shape the relationship between empathy and manipulating behaviours (Burgmer et al., 2021; Konrath et al., 2013). Looking at these findings, the question arises whether narcissism could influence the relation between empathy and the acceptance of gaslighting behaviours. The rationale being that individuals with higher scores on narcissism and empathy might be more accepting of gaslighting behaviours than people with high empathy and low narcissism. Therefore, the possibility arises to confirm the findings by March et al. (2023), regarding the effect of narcissism on gaslighting, while also investigating a potential moderation effect between narcissism and the relationship between empathy and gaslighting. In conclusion, this paper will answer the following questions:

1) Is empathy negatively associated with the acceptance of gaslighting?
2) Is the relation between empathy on the acceptance of gaslighting behaviours moderated by scores on narcissism?

Methods

Study Design

The present study concerns a cross-sectional online survey study among 147 individuals of mainly Dutch and German nationality performed in the course of the bachelor thesis of 3 psychology students. The current study investigates the relationship between empathy and acceptance towards gaslighting behaviours, with narcissism as a potential moderator for this relationship. Participants were recruited using non-probability, convenience sampling through the University of Twente’s SONA system. Since this would lead to a disproportionate amount of young and highly educated university (psychology) students in the sample, participants were also recruited from the personal networks of the researchers to further diversify the sample and increase the number of participants.

Material and Instruments

The questionnaire consisted of eight distinct blocks, each measuring different constructs using previously validated questionnaires. The blocks encompassed the following

constructs: Gaslighting acceptance, attachment styles, emotional intelligence, self-esteem, desirability of control, empathy, narcissism, and alexithymia. For this study in particular *gaslighting acceptance*, *narcissism* and *empathy* were of relevance. Before filling in the previously mentioned tests, the participants were asked about their demographics, field of study or occupation, and whether they were university students. The data for this study was collected using the online survey platform Qualtrics (www.qualtrics.com), allowing participants to partake in the survey using their own electronic devices. The questionnaire was administered in English and in total the survey consisted of 167 items.

Gaslighting acceptance was measured with the scale developed by March et al. (2023). The scale was validated in a sample of 315 Australians aged between 18 and 82, but March et al. (2023) recommend further validation of the questionnaire with different samples. The scale demonstrates good internal reliability with a Cronbach's alpha of .97. Overall, the questionnaire consists of 10 statements about scenarios between two people, wherein person A tries to manipulate person B. Participants have to indicate how much they agree with the statements on a Likert scale from 1 (unacceptable) to 7 (acceptable). The mean score was calculated, with higher scores indicating higher acceptance towards gaslighting behaviours.

The participants' score on narcissistic tendencies was assessed using the Brief Pathological Narcissism Inventory (B-PNI). This questionnaire was developed and validated in the study by Schoenleber et al. (2015) and consists of 28 items investigating vulnerable and grandiose narcissism. The items include statements about attitudes and behaviours, and participants must indicate to which degree they identify with the statement on a 6-point Likert scale. Schoenleber et al. (2015) collected responses on the B-PNI and other narcissism inventories, as well as other related scales, from a heterogeneous sample of 3851 participants. The results showed a good internal consistency for both vulnerability (min. $\alpha = .93$) and grandiosity (min. $\alpha = .83$). Both subscales were used in this study. The mean scores were calculated, with higher scores being indicative of more narcissistic individuals.

For the measurement of participants' empathy, the Toronto Empathy Questionnaire (TEQ) was used. Since it has been validated and proven reliable and due to its short length, it has been selected for this study. The TEQ consists of 16 items in the form of short statements regarding emotional responses to a given situation. Participants are asked to indicate how frequently they experience these emotional responses on a scale from 0 (never) to 4 (always). The questionnaire was developed in a study by Spreng et al. (2009), wherein the authors collected a total of 142 items and eliminated those with low item-remainder coefficients. The remaining 16 items were included in the TEQ which on further investigation showed a good internal consistency with Cronbach's $\alpha = .85$. Normally, the TEQ has a 5-point likert scale, but in the present study, inadvertently, a 4-point likert

scale was used. To compute mean scores, negative items were reverse coded. More empathetic individuals are thought to have higher mean scores on the scale.

Procedure

Prior to the implementation of the study, ethical approval was obtained by the Ethics committee at the University of Twente (240367). After getting approval, the study was added to the online platform SONA through which participants were recruited. Additionally, the convenience sample of participants was also recruited through social media platforms. Before filling in the survey, the respondents were informed about their rights and the procedure of the study. After obtaining all the information, the participants gave their informed consent to participate in the study. It should be noted that the information given did not explicitly mention that the study measured acceptance of gaslighting, to prevent any potential bias, and rather gave a general description. Once the survey was completed, the participants who were recruited through the SONA platform received credits for participating in the study.

Data Analysis

All analyses were performed in RStudio version 2021.09.0. After data collection, the data was screened for rows containing duplicate IP addresses and of rows with missing value. These rows (i.e. participants) were excluded from the final dataset. Next, descriptive analyses were conducted to determine sample characteristics. Afterwards, items of the questionnaires (i.e. acceptance of gaslighting, empathy, narcissism) were reverse scored if needed and the total scores were calculated. Using these total scores, a Pearson correlation matrix was constructed to test for univariate correlations between individual variables. In addition to that, histograms were plotted for the total scores on each questionnaire to check for the normal distribution of scores.

For the main analysis, the PROCESS macro for RStudio was used to run a moderation model (model 1) with 10000 bootstrap iterations (Hayes, 2018). This approach was chosen over the approach using interaction terms in regression models proposed by Baron and Kenny (1986). The results of Hayes (2018) method are generally considered more robust, because it makes use of bootstrapping. For the model, acceptance of gaslighting was considered the independent variable, while empathy was the dependent variable and narcissism the moderator of the relation between the two. The significance of any (main or interaction) relation is established when the bootstrap confidence intervals do not contain 0. For three models (generalised narcissism, vulnerability or grandiosity as the moderator), the five assumptions of moderation models were tested, namely: linearity, independence, homoscedasticity, normality of residuals, and multicollinearity. Linearity was tested by visually inspecting a plot of the residuals against the fitted values. Similarly, the normality of residuals was investigated by visual assessment of the Q-Q plot, but also by

performing a Saphiro-Wilk test and a Kolmogorov-Smirnov test. Homoscedasticity was evaluated using the Breusch-Pagan test, and the non-constant variant test. A Durbin-Watson test was used to assess independence, and the Variance Inflation Factor was used to test the multicollinearity.

Next, a sensitivity moderation analysis was run, including age and gender as covariates, to check for an increase in explained variance and deviating results. March et al. (2023) already showed a significant relation of gender on gaslighting acceptance scores; therefore it was considered reasonable to run a model accounting for this. It might also be the case that older individuals are less accepting of gaslighting due to experience. They might be more able to identify manipulating behaviour and thus be less accepting towards the statements in the acceptance of gaslighting questionnaire. Simple slopes plots of the moderation effects were generated for the main analysis, as well as for the sensitivity analysis. PROCESS provides the option to include data that can be used for visualisation in the output of the analysis. This output was then used to manually create a plot using the *ggplot* function of the *ggplot2* package.

Finally, the psychometric properties of the acceptance of gaslighting questionnaire were reassessed. For this purpose, a strict unidimensional (1-factorial) model was tested with confirmatory factor analysis (CFA) using the *lavaan* package in RStudio. This was done because the questionnaire used to measure acceptance of gaslighting is a relatively new one. Marchet al. (2023) just recently composed it in their study and called for further validation of the scale in different samples. To assess the model fit, measurements and cutoff points were taken from Hu and Bentler (1999). According to this study, a non-significant ($p > .05$) chi-squared value would indicate good model fit. The root mean square error of approximation (RMSEA) has the following cutoff points: RMSEA < 0.06 indicates a good fit, and 0.06 ≤ RMSEA < 0.08 indicates acceptable fit. Next, a standardised root mean square residual (SRMR) smaller than 0.08 indicates a good fit. The comparative fit index (CFI) has the cutoff points as follows: CFI ≥ 0.95 indicates a good fit; 0.90 ≤ CFI < 0.95 indicates acceptable fit. And finally, the cutoff points for the Tucker-Lewis index (TLI): TLI ≥ 0.95 indicates a good fit; 0.90 ≤ TLI < 0.95 indicates acceptable fit. In terms of measurement quality, standardised factor loadings above .40 are taken to be sufficient, according to Van Zyl and Ten Klooster (2022).

Results

Participants

In total 147 people started with the survey. After cleaning the data, the study sample consisted of 96 participants whose ages ranged from 18 to 61 years, with a mean age of 24.99 years (SD = 8.12). The majority of participants were female (n = 66), with 29 participants identifying as male and one participant identifying as another gender. The

distribution of nationalities indicated was 55 Germans participants, 16 Dutch, and 25 people from other nationalities. Regarding educational status, the majority of participants (*n* = 80) reported currently studying, while 16 indicated they were not currently studying. Within the sample, 54 of participants were studying psychology at the time of the study.

Distribution of Scale Scores

The observed scores on the acceptance of gaslighting questionnaire ranged from 10 to 54, with a mean score of 18.24 (SD = 8.22). The distribution of scores was as follows: the first quartile (Q1) was 12, the median (Q2) was 17, and the third quartile (Q3) was 21. The variance of the scores was 67.59. The mean score for males on the acceptance of gaslighting questionnaire was 20.6, while the mean score for females was 17.1. Total scores showed a heavy right skew in responses on the acceptance of gaslighting questionnaire, as can be seen in Figure 1. For the other questionnaires the skew was less intense, but for both still present, as can be seen in Appendix B. Scores on the TEQ were skewed slightly to the left and scores on the B-PNI slightly to the right.

Univariate Correlations Between the Constructs

As part of these preliminary analyses, a correlation analysis was conducted to examine the relationship between empathy, narcissism, and scores on the acceptance of gaslighting questionnaire. The results of this analysis are summarised in Table 1. As expected, empathy scores were negatively correlated ($p < .001$) with gaslighting acceptance. Narcissism scores were positively associated with gaslighting ($p = .058$), and negatively, but not significantly, correlated with empathy ($p = .86$).

Table 1

Pearson Correlations between the Three Main Variables

	Gaslighting Acceptance	Empathy
Empathy	-0.35***	1.000
Narcissism	0.19	-0.02

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

Unadjusted Moderation Analyses

Following the univariate correlations, three moderation analyses were run with *generalised narcissism, vulnerable narcissism, and grandiose narcissism* as the respective moderators. The models examined the influence of narcissism on the relationship between empathy and gaslighting acceptance, while considering age and gender as covariates. The results of these moderation analyses can be found in Table 2.

In summary, the expected direct negative relation of empathy on acceptance of gaslighting scores was found to be significant in all moderation models, as evidenced by zero not being included in the 95% bootstrap confidence intervals. Additionally, there was a significant positive relation of generalised narcissism on the acceptance of gaslighting. Interestingly, when looking at the subscales, grandiose narcissism had a comparatively small and insignificant positive relation, while vulnerable narcissism on its own had an even stronger, significant positive relation.

The interaction effects of generalised narcissism and its two subscales were not substantial and insignificant in all three models, indicating that narcissism did not moderate the relation between empathy and gaslighting acceptance in any of the models.

Figure 1

Histogram of Scores on the Acceptance of Gaslighting Questionnaire

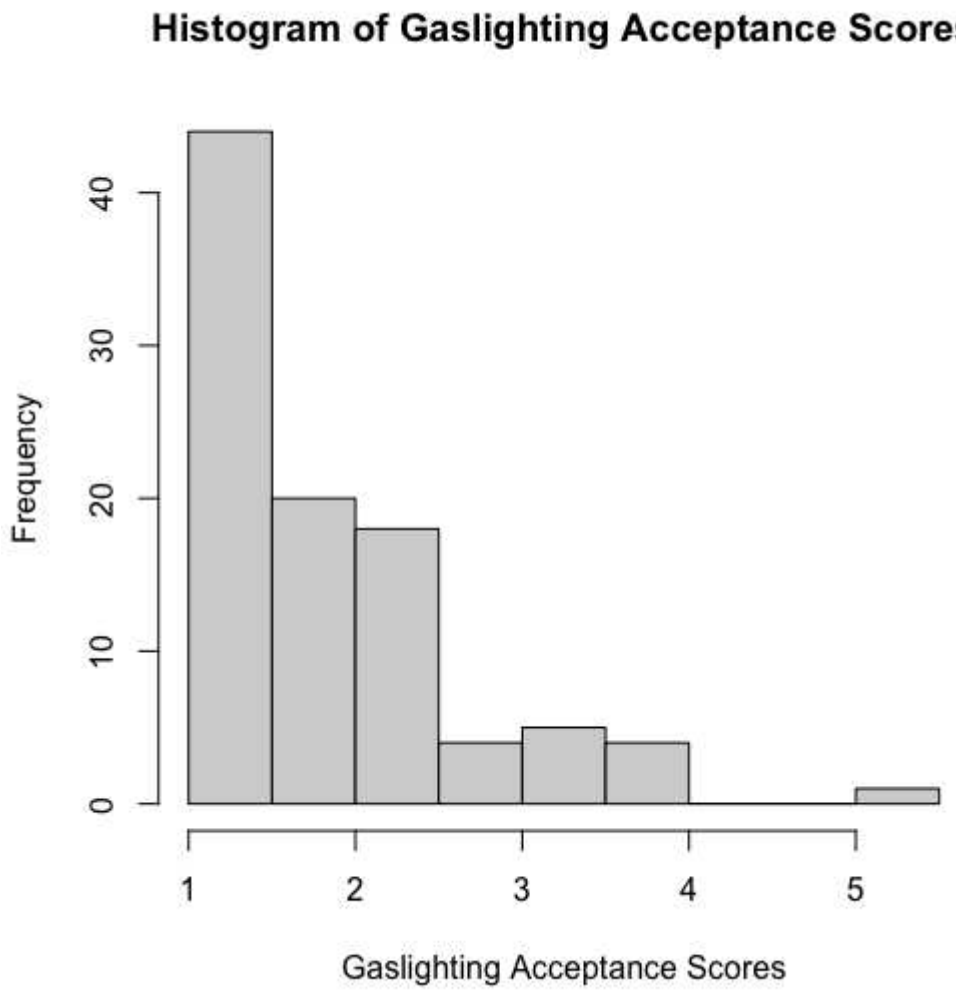


Table 2

Bootstrap Results of the unadjusted Moderation Analyses

	Coefficient	Mean	SE	Confidence Interval Low	Confidence Interval High
Using Generalised Narcissism as the Moderator					
Constant	1.82	1.83	0.08	1.68	1.99
Empathy	-0.83	-0.82	0.31	-1.42	-0.16
Narcissism	0.23	0.25	0.11	0.04	0.49
Interaction	0.13	-0.02	0.45	-1.10	0.64
Using Vulnerable Narcissism as the Moderator					
Constant	1.82	1.83	0.08	1.68	1.99
Empathy	-0.90	-0.89	0.32	-1.48	-0.24
Vulnerable Narcissism	0.27	0.29	0.11	0.09	0.52
Interaction	0.29	0.17	0.41	-0.78	0.84
Using Grandiose Narcissism as the Moderator					
Constant	1.82	1.82	0.08	1.67	1.98
Empathy	-0.81	-0.79	0.31	-1.40	-0.13
Grandiose Narcissism	0.12	0.14	0.10	-0.05	0.33
Interaction	-0.03	-0.14	0.37	-1.02	0.45

Note. Generalised Narcissism: R = .40, F(3, 92) = 5.86, p = .001, explaining 16.04% (R-squared) of the variance with an MSE of 58.60. Vulnerable Narcissism: R = 0.43, F(3, 92) = 7.02, p < .001, explaining 18.62% (R-squared) of the variance with an MSE of 56.80. Grandiose Narcissism: R = .37, F(3, 92) = 4.95, p = .003, explaining 13.91% (R-squared) of the variance with an MSE of 60.09.

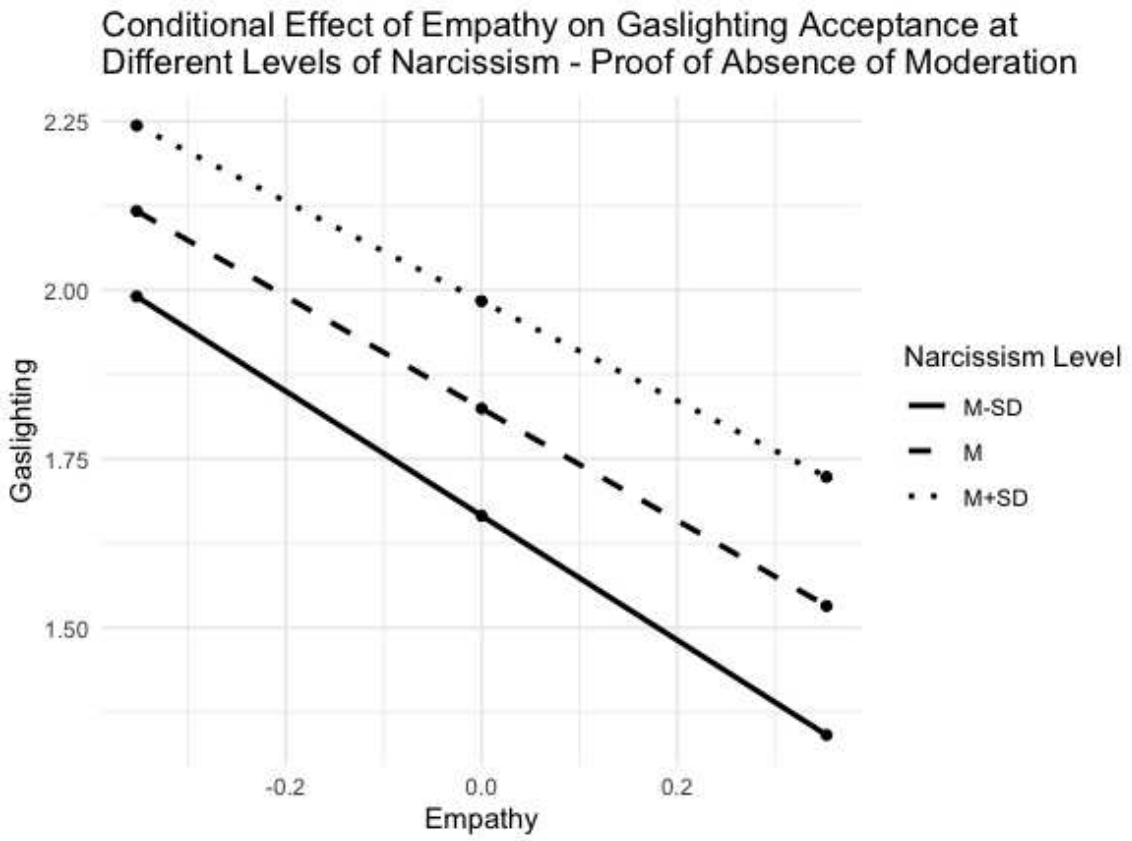
Age and Gender Adjusted Moderation Analyses

In the sensitivity analysis, age too had insignificant relation on acceptance of gaslighting in all three models. Gender had a significant relation on acceptance of gaslighting when included in the model for grandiose narcissism, with males having higher scores on average. The relation was equally as large for the other models, but for them, the bootstrap intervals included 0.

As with the unadjusted moderation models, all interaction terms for narcissism were insignificant. The absence of a moderating effect of narcissism on the relation between empathy and acceptance of gaslighting is illustrated in Figure 2. This figure shows the very similar negative associations between empathy and gaslighting acceptance for people with lower, medium and higher levels of generalised narcissism. The numerical results of the sensitivity analysis are summarised in Appendix C.

Figure 2

Simple Slopes Plot of Moderation Analysis Results



Model Fit and Measurement Quality of the Acceptance of Gaslighting Questionnaire

The results of the CFA of the acceptance of gaslighting questionnaire indicated a less bad fit for a strict unidimensional model based on the model fit indices ($\chi^2(35) = 94.76$, $p < .001$, CFI = 0.86, TLI = 0.82, and RMSEA = 0.13 (90% CI = [0.10; 0.17]), SRMR = 0.07). Despite the bad fit indices, all standardised factor loadings were substantial and statistically significant ($p < .001$), ranging from 0.56 to 0.76, indicating adequate measurement quality (see Table 3). Cronbach's alpha for the scale in this sample was 0.89.

Table 3

Factor Loadings of Items in the Acceptance of Gaslighting Questionnaire, Assuming One Underlying Factor

Item	Statement	Standardised Factor Loading
1	Person A accuses Person B of lying, even when Person A knows that they are the one who is lying	0.71
2	Person A tells Person B that they are wrong, even when Person A knows that what Person B is saying is true	0.69
3	Person A accuses Person B of being paranoid, even if Person A knows that Person B's suspicions are well-founded	0.68
4	Person A tries to make Person B question their sanity	0.57
5	Person A says anything to Person B if it means that they will get their way	0.56
6	Person A lashes out at Person B whenever Person B says something that contradicts Person A's version of events	0.71
7	Person A never admits to doing anything wrong, even when Person B has proof that Person A did do something wrong	0.68
8	Person A says Person B has a bad memory if Person B catches Person A telling a lie	0.57
9	Person A makes Person B question their decision-making abilities, if it means Person A gets to be the one to make decisions in the relationship	0.74
10	Person A lies to Person B just to see if Person B will believe them	0.76

Discussion

This study set out to investigate the relationship between empathy and the acceptance of gaslighting and considered the role of narcissism as a potential moderator of this relationship. Another goal was the confirmation of the results from the study by March et al. (2023), regarding the positive relationship between narcissism and gaslighting acceptance. Additionally, the validity and reliability of the very recently developed gaslighting acceptance questionnaire were further investigated in a new sample.

In regards to the first research question, the relationship between empathy and the acceptance of gaslighting, it can be concluded that there was a significant negative relationship between the two constructs. Individuals with higher empathy tended to score

lower on the gaslighting acceptance questionnaire. A possible explanation for this expected finding could be that a greater capacity to understand other people's worldview and their emotion makes people less inclined to find it acceptable to make another person doubt their own reality. As Sodoma (2022) explained, empathy relates to the capacity of individuals to understand another person's cognitive and emotional experience. Heym et al. (2021), on the other hand point out that so called dark empaths make use of exactly this skill to manipulate people more effectively. These people could be more aware of another person's subjective experience, and then misuse this understanding for personal gain. It could be argued that those individuals would have to have a more explicit malicious intent and perhaps score higher on dark trait traits. March et al. (2023) investigated the influence of dark tetrad traits on acceptance of gaslighting, and indeed found a positive relationship.

These findings were also replicated in the current study. Specifically, March et al. (2023) found the highest positive relationship to be between vulnerable narcissism and gaslighting acceptance. This was confirmed in the present findings as, the relationship between generalised narcissism and acceptance of gaslighting was found to be less substantial, and less significant than the relationship between vulnerable narcissism and gaslighting acceptance. This is probably due to the finding that the relationship between grandiose narcissism and acceptance of gaslighting was found to be insubstantial and insignificant. It is likely that the inclusion of grandiose narcissism in the generalised narcissism score diminished both effect size and significance of the relation of generalised narcissism. March et al. (2023) stated that individuals high in vulnerable narcissism possess a number of characteristics that make them more accepting towards gaslighting behaviours. Mainly, these individuals might see gaslighting strategies as a way of gaining control over their partner and their opinions, to minimise the threat of negative feedback (March et al., 2023). Green et al. (2020) also stated that individuals with high vulnerable narcissism scores were more likely to perpetrate psychological abuse, which gaslighting is a form of. All this might support the direct positive relation of narcissism on the acceptance of gaslighting scores shown in the present study.

As for the interaction effect of narcissism and empathy on acceptance of gaslighting scores, the results showed no signs of any significant interaction effect. For all three narcissism scores, the interaction effects were small and insignificant. The absence of any interaction effects can also be observed in the simple slopes figure, which showed no effect of different levels of narcissism on the relation between empathy and acceptance of gaslighting. It has to be mentioned, however, that the sample size for this study was rather small to properly test for an interaction effect, as it was powered to detect a moderate main (direct) relation only. Another reason for the absence of a moderating effect could be that empathy and narcissism are too broad constructs. Konrath et al. (2013) for example,

mentioned that specifically emotion recognition abilities are correlated with exploitative narcissism. This leaves open the possibility for an interaction between related constructs of empathy and narcissism.

The sensitivity analysis, which included age and gender as covariates, also showed no sign of an interaction effect (see Appendix C). The main difference was in the relation between grandiose narcissism and acceptance of gaslighting. In the adjusted models, the direct relation seemed to become more substantial and significant. Interestingly, when including age and gender as covariates, gender demonstrated a substantial direct relation in the analysis, indicating that men tended to be more accepting of gaslighting behaviours. Effect sizes range from -0.3 to -0.43 for females, but the confidence intervals, in all cases, still contain 0 indicating that this gender relation was not significant. These findings are partly in line with the findings of March et al. (2023). In their study they also found that gender had a large, but in their case significant, relation with gaslighting acceptance scores. However, they also found a rather large interaction effect of age on the relation between dark tetrad traits, and specifically vulnerable narcissism. The effect size of gender in this study certainly echoes the results from March et al. (2023), but the results lack significance, again perhaps due to a lack of statistical power as a result of the relative underrepresentation of men in this sample. Furthermore, this study did not investigate the interaction effect between gender and narcissism.

This study additionally investigated the factorial validity and reliability of the acceptance of gaslighting questionnaire in a new sample. March et al. (2023) advised since the questionnaire was recently developed and needs to be investigated with multiple different samples. Conveniently, this analysis also enabled the investigation of the quality of the scores on the acceptance of gaslighting questionnaire. Overall, the analysis showed that none of the fit indices for a strict unidimensional model reached sufficient cutoff points. However, as has already been pointed out by Schermelleh-Engel et al. (2003), these indices do not necessarily mean that the gaslighting acceptance questionnaire fails to measure a unidimensional trait. Sufficient model fit would indicate that the model is plausible. In the same vein, Perry et al. (2015) mentioned that the cutoff points proposed by Hu and Bentler (1999) are, generally speaking, often too strict. Furthermore, the factor loadings all succeed the .04 cutoff point for newly developed items, as proposed by Van Zyl and Ten Klooster (2022). For the present study it was decided that the questionnaire sufficiently measures the latent construct of gaslighting acceptance. However, future studies should further examine the factor structure of the acceptance of gaslighting questionnaire in different samples.

In regard to the reliability of the gaslighting acceptance questionnaire and the other scales that were used, it was found that all of them displayed sufficient reliability. All of the Cronbach’s Alphas succeed 0.8, which is generally deemed a good fit (Cicchetti, 1994). It

cannot go unmentioned that Cronbach's Alpha has received its fair share of critique. Cho and Kim (2015) address all sorts of misconceptions about the Cronbach Alpha, for example whether “a high value of alpha is indicative of internal consistency” (p. 207). It does however show if participants who answered high on one item of the scale were likely to answer high on other items as well. And since Cronbach’s Alpha has been used across the field for a long time now, and is still considered the standard reliability measure, it was considered to suffice for the purpose of this study. Caution is still advised in interpreting the presented results, as there was no other reliability measure included.

In terms of validity of the main study findings, a strength of the study is that it used previously validated scales for all constructs of interest. The B-PNI and the TEQ have long been established in the field for clinical and research purposes. This, however, cannot be said for the acceptance of gaslighting questionnaire. The present analysis provided some evidence for the validity of the scores on this scale, but there needs to be more research and practical application before its validity is properly established. Furthermore, although the TEQ is well established by now, a 4-point likert scale was used accidentally, which compromises its reliability. In practical terms, the variability of scores on the questionnaire is likely to be dampened. This means that relations that were found may appear less significant than they actually are, which also reduces the chance of a type 1 error. It must also be mentioned that the present study did not account for social desirability effects despite the expected risk regarding acceptance of gaslighting and narcissism. It might well be that participants underreported their respective scores on acceptance of gaslighting and narcissism.

In terms of external validity, however, it needs to be pointed out that after cleaning the data there were 96 participants left which is a rather small sample size for the analyses that were executed with respect to the moderation analysis. Additionally, a majority of the sample consisted of university students, specifically psychology students. With this, the study suffers from the same sample bias as most psychological studies (Heinrich et al., 2010). Furthermore, the sample consisted mostly of female identifying individuals between the ages of 21 and 25, which is not representative of the general population. Nonetheless, the results from this study show substantial evidence for an association between empathy and acceptance of gaslighting scores, which is consistent with the literature.

This study therefore contributed to the further understanding of gaslighting acceptance and its antecedents. In addition, the results of March et al. (2023), regarding the relationship between (vulnerable) narcissism and acceptance of gaslighting could be replicated in this sample. Finally, the reliability and validity of the acceptance of gaslighting questionnaire were partly confirmed in our findings. With the caveats of the present study it is recommended to investigate the relation of empathy on gaslighting acceptance in a larger

sample. It is also recommended to include the full 5-point likert scale for the TEQ in future studies, and to include a scale to measure social desirability in the sample. Furthermore, with a substantially larger sample size, a potential moderation effect of narcissism on the relation between empathy and narcissism can be ruled out or substantiated. It is also recommended to include age and gender as covariates in the models investigating these relationships, as well as to check for interaction effects, especially in regards to gender. The prospect of investigating the constituents of empathy, in addition to related constructs, and their relation to acceptance of gaslighting, also seems a promising direction for further research.

References

APA Dictionary of Psychology. (n.d.). https://dictionary.apa.org/gaslight

Awang, Z. (2014). *A Handbook onStructural Equation Models* (2nd ed.). MPWS Rich Resources.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173. https://doi.org/10.1037/0022-3514.51.6.1173

Bibeau, M., Dionne, F., & Leblanc, J. (2015). Can compassion meditation contribute to the development of psychotherapists' empathy? a review. *Mindfulness*, *7*(1), 255–263. https://doi.org/10.1007/s12671-015-0439-y

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well known but poorly understood. *Organizational research methods*, *18*(2), 207-230. https://doi.org/10.1177/1094428114555994

Cho, H., & Lee, D. (2023). Effects of affective and cognitive empathy on compassion fatigue: mediated moderation effects of emotion regulation capability. *Personality and Individual Differences, 211*. https://doi.org/10.1016/j.paid.2023.112264

Crossman, K. A., Hardesty, J. L., & Raffaelli, M. (2015). "He Could Scare Me Without Laying a Hand on Me". *Violence Against Women*, *22*(4), 454-473. https://doi.org/10.1177/1077801215604744

Dor-Ziderman, Y., Cohen, D., Levit-Binnun, N., & Golland, Y. (2021). Synchrony with distress in affective empathy and compassion. *Psychophysiology, 58*(10), 13889. https://doi.org/10.1111/psyp.13889

Ferrer-Perez, V. A., Bosch-Fiol, E., Ferreiro-Basurto, V., Delgado-Alvarez, C., & Sánchez-Prada, A. (2020). Comparing Implicit and Explicit Attitudes Toward Intimate Partner Violence Against Women. *Frontiers in psychology*, *11*, 2147. https://doi.org/10.3389/fpsyg.2020.02147

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, *33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Heym, N., Kibowski, F., Bloxsom, C. A. J., Blanchard, A., Harper, A., Wallace, L., Firth, J., & Sumich, A. (2021). The dark empath: characterising dark traits in the presence of

empathy. *Personality and Individual Differences*, *169*. https://doi.org/10.1016/j.paid.2020.110172

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. https://doi.org/10.1080/10705519909540118

Israelashvili, J., Sauter, D., & Fischer, A. (2020). Two facets of affective empathy: concern and distress have opposite relationships to emotion recognition. *Cognition & Emotion*, *34*(6), 1112–1122. https://doi.org/10.1080/02699931.2020.1724893

Karnaze, M. M., Bellettiere, J., & Bloss, C. S. (2022). Association of compassion and empathy with prosocial health behaviors and attitudes in a pandemic. *Plos One*, *17*(7), 0271829. https://doi.org/10.1371/journal.pone.0271829

Kappelmayer, M., Czar, A., Tresca, M., D'Adamo, P., & Lozada, M. (2023). A school intervention promotes compassion, empathy and social relationships in children. *School Psychology International*, *44*(5), 515–532.

March, E., Kay, C. S., Dinić, B. M., Wagstaff, D., Grabovac, B., & Jonason, P. K. (2023). "It's All in Your Head": Personality Traits and Gaslighting Tactics in Intimate Relationships. *Journal of Family Violence*. https://doi.org/10.1007/s10896-023-00582-y

Miller, J. T., & Verhaeghen, P. (2022). Mind full of kindness: self-awareness, self-regulation, and self-transcendence as vehicles for compassion. *Bmc Psychology*, *10*. https://doi.org/10.1186/s40359-022-00888-4

Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing Model Fit: Caveats and Recommendations for Confirmatory Factor Analysis and Exploratory Structural Equation Modeling. *Measurement in Physical Education and Exercise Science*, *19*(1), 12–21. https://doi.org/10.1080/1091367X.2014.952370

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, *8*(2), 23-74.

Sodoma, K. A. (2022). Emotional gaslighting and affective empathy. *International Journal of Philosophical Studies*, *30*(3), 320–338. https://doi.org/10.1080/09672559.2022.2121894

Stark, C. A. (2019). Gaslighting, Misogyny, and Psychological Oppression. *The Monist, 102*(2), 221–235. https://doi.org/10.1093/monist/onz007

Stevens, F., & Taber, K. (2021). The neuroscience of empathy and compassion in pro-social behavior. *Neuropsychologia*, *159*.

Van Zyl, L. E., & Ten Klooster, P. M. (2022). Exploratory structural equation modeling: Practical guidelines and tutorial with a convenient online tool for Mplus. *Frontiers in Psychiatry*, *12*, 795672. https://doi.org/10.3389/fpsyt.2021.795672

Appendix A

RStudio Code for Data Analysis

```
# Load necessary libraries
library(broom)
library(car)
library(Hmisc)
library(readr)
library(dplyr)
library(ggplot2)
library(lavaan)
library(ltm)
library(lmtest)
library(tidyr)
library(tidyverse)



# Load the dataset
raw_data <- read_csv("gaslighting_RAW.csv")


#Select relevant variables
dataset <- raw_data %>% select("IPAddress", "Q1":"Q8_2_TEXT", "Q56_1":"Q65_1",
"Q95":"Q119", "Q122":"Q149")


dataset <- dataset[-c(1, 2), ]
dataset <- as.data.frame(dataset)


#Formatting variables approriately
numeric_var <- c(2:7, 9, 10, 12:65)
```

```
dataset <- dataset %>%
  mutate_at(vars(numeric_var), as.numeric)

##Checking for duplicate IP-Addresses
# Extract the IP address column
ip_addresses <- dataset$IPAddress
# Find duplicate IP addresses
duplicate_ips <- ip_addresses[duplicated(ip_addresses)]
# Print duplicate IP addresses
print(duplicate_ips)

# Filter out rows with missing values and duplicte IPs
# Create a logical condition to identify rows with duplicate IP addresses
duplicate_rows <- duplicated(dataset$IPAddress) | duplicated(dataset$IPAddress,
fromLast = TRUE)
# Subset the dataset to exclude rows with duplicate IP addresses
dataset <- dataset[!duplicate_rows, ]
#Exclude rows with missing values
quest_results <- c(12:65)
results <- dataset[complete.cases(dataset[, quest_results]), ]

##Rename Columns
results <- results %>%
  rename(Age = Q4, Gender = Q5)

##Recode Gender
results$Gender <- factor(results$Gender, levels = c(1, 2, 3), labels = c("Male",
"Female", "Other"))
```

```
##Demographics
#Age
summary(results$Age)
#Nationality
table(results$Q6) %>% print()
results %>%
  count(Q6_3_TEXT, sort = TRUE) %>%
  print()
#Gender
table(results$Gender) %>% print()
#Study / Occupation
table(results$Q7) %>% print()
table(results$Q8) %>% print()
results %>%
  count(Q8_2_TEXT, sort = TRUE) %>%
  print()

##Defining items belonging to questionnaires and (relevant) subscales
gaslighting <- c(12:21)
empathy <- c(22:37)
narcissism <- c(38:65)
vul_narc <- c(41,43,38,48,61,56,47,49,63,54,62,50)
g_narc <- c(58,53,39,55,65,40,52,64,57,51,46,44,59,42,45,60)

##Recoding variables to fit the likert scale
likert_ajust <- c(22:65)
results[, likert_ajust] <- results[, likert_ajust] - 1
```

```
##Reverse coding variables
empathy_reverse <- c(23, 25, 28, 31, 32, 33,35, 36)
results[, empathy_reverse] <- 3 - results[, empathy_reverse]

##Computing mean scores
results$gaslighting <- rowMeans(results[, gaslighting])
results$empathy <- rowMeans(results[, empathy])
results$narc <- rowMeans(results[, narcissism])
results$vul_narc <- rowMeans(results[, vul_narc])
results$g_narc <- rowMeans(results[, g_narc])


#Correlation Matrix
correlation_results <- results %>% select(gaslighting, empathy, narc)
univar_corr <- rcorr(as.matrix(correlation_results))
univar_corr$r
univar_corr$P

##Cronbach's Alpha
cronbach.alpha(results[gaslighting])
cronbach.alpha(results[empathy])
cronbach.alpha(results[narcissism])
cronbach.alpha(results[vul_narc])
cronbach.alpha(results[g_narc])

##Descriptive statistics
#Descriptives
```

```
summary(results$gaslighting)
sd(results$gaslighting)
var(results$gaslighting)
table(results$gaslighting)

# Histogram example
hist(results$gaslighting, main="Histogram of Gaslighting Acceptance Scores", xlab =
"Gaslighting Acceptance Scores")
hist(results$empathy, main="Histogram of Empathy Scores", xlab = "Empathy
Scores")
hist(results$narc, main="Histogram of Narcissism Scores", xlab = "Narcissism
Scores")

##Mediation model
#Hayes PROCESS analysis
process(data = results, y = "gaslighting",
      x = "empathy", w ="narc",
      model = 1, center = 2, moments = 1,
      plot = 1, jn = 1,
      modelbt = 1, boot = 10000, seed = 654321)
process(data = results, y = "gaslighting",
      x = "empathy", w ="vul_narc",
      model = 1, center = 2, moments = 1,
      plot = 1, jn = 1,
      modelbt = 1, boot = 10000, seed = 654321)
process(data = results, y = "gaslighting",
      x = "empathy", w ="g_narc",
      model = 1, center = 2, moments = 1,
```

```
        plot = 1, jn = 1,
        modelbt = 1, boot = 10000, seed = 654321)

#Sensitivity Analyses
results$Gender <- as.numeric(factor(results$Gender, levels = c("Male", "Female",
"Other"), labels = c(1, 2, 3)))

process(data = results, y = "gaslighting",
        x = "empathy", w ="narc",
        model = 1, center = 2, moments = 1,
        plot = 1, jn = 1, cov = c("Age", "Gender"),
        modelbt = 1, boot = 10000, seed = 654321)
process(data = results, y = "gaslighting",
        x = "empathy", w ="vul_narc",
        model = 1, center = 2, moments = 1,
        plot = 1, jn = 1, cov = c("Age", "Gender"),
        modelbt = 1, boot = 10000, seed = 654321)
process(data = results, y = "gaslighting",
        x = "empathy", w ="g_narc",
        model = 1, center = 2, moments = 1,
        plot = 1, jn = 1, cov = c("Age", "Gender"),
        modelbt = 1, boot = 10000, seed = 654321)

##Checking Assumptions of the Moderation Model
#Recreating the Model
# Fit the linear model
asmpt_test <- lm(gaslighting ~ empathy * narc, data = results)
# Linearity
```

```
plot(asmpt_test$fitted.values, asmpt_test$residuals)
abline(h = 0, col = "red")
# Independence
dwtest(asmpt_test)
# Homoscedasticity
bptest(asmpt_test)
ncvTest(asmpt_test)
# Normality of Residuals
qqnorm(asmpt_test$residuals)
qqline(asmpt_test$residuals, col = "red")
shapiro.test(asmpt_test$residuals)
ks.test(asmpt_test$residuals, "pnorm", mean = mean(asmpt_test$residuals), sd =
sd(asmpt_test$residuals))
# Multicollinearity
vif(asmpt_test, type = 'predictor')

##Plots
#Output of the moderation analysis
moderation_plot <- data.frame(
  empathy = c(-0.3527, 0.0000, 0.3527, -0.3527, 0.0000, 0.3527, -0.3527, 0.0000,
0.3527),
  narc = c(-0.6951, -0.6951, -0.6951, 0.0000, 0.0000, 0.0000, 0.6951, 0.6951,
0.6951),
  gaslighting = c(1.9904, 1.6657, 1.3410, 2.1170, 1.8245, 1.5321, 2.2435, 1.9834,
1.7232)
)

# Plot the interaction effect
```

```
custom_labels <- c("-0.6951" = "M-SD", "0" = "M", "0.6951" = "M+SD")


# Plot the interaction effect
ggplot(moderation_plot, aes(x = empathy, y = gaslighting, linetype = as.factor(narc)))
+
  geom_line(linewidth = 1) +
  geom_point() +
  scale_linetype_manual(values = c("solid", "dashed", "dotted"), labels =
custom_labels) +
  labs(
    title =
"Conditional Effect of Empathy on Gaslighting Acceptance at
Different Levels of Narcissism - Proof of Absence of Moderation",
    x = "Empathy",
    y = "Gaslighting",
    linetype = "Narcissism Level"
  ) +
  theme_minimal()


#Plot of the moderation analysis with sensitivity covariates
sens_moderation_plot <- data.frame(
  empathy = c(-0.3535, 0.0000, 0.3535, -0.3535, 0.0000, 0.3535, -0.3535, 0.0000,
0.3535),
  narc = c(-0.6996, -0.6996, -0.6996, 0.0000, 0.0000, 0.0000, 0.6996, 0.6996,
0.6996),
  gaslighting = c(1.9072, 1.6142, 1.3213, 2.0900, 1.8129, 1.5358, 2.2728, 2.0115,
1.7503)
)
```

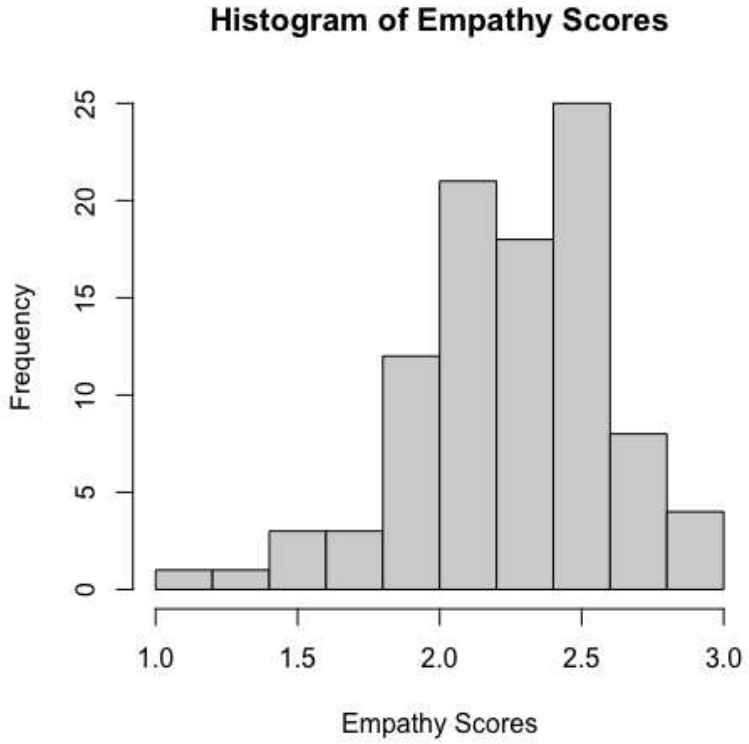
```
sens_custom_labels <- c("-0.6996" = "M-SD", "0" = "M", "0.6996" = "M+SD")


ggplot(sens_moderation_plot, aes(x = empathy, y = gaslighting, linetype =
as.factor(narc))) +
  geom_line(linewidth = 1) +
  geom_point() +
  scale_linetype_manual(values = c("solid", "dashed", "dotted"), labels =
(sens_custom_labels)) +
  labs(
    title = "Proof of Absence of Moderation
With Sensitivity Measures",
    x = "Empathy",
    y = "Gaslighting",
    linetype = "Narcissism Level"
  ) +
  theme_minimal()


##CFA of the Gaslighting Questionnaire
acceptance_gaslighting <- '
Acceptance_Gaslighting =~ Q56_1 + Q57_1 + Q58_1 + Q59_1 + Q60_1 + Q61_1 +
Q62_1 + Q63_1 + Q64_1 + Q65_1
'
fit <- cfa(acceptance_gaslighting, data = results)
summary(fit, fit.measures = TRUE, standardized = TRUE)
standardizedsolution(fit)
```

Appendix B

Histograms of Test Scores on Independent Constructs/Variables



Appendix C

Bootstrap Results of the Moderation Analyses With Covariates for Sensitivity

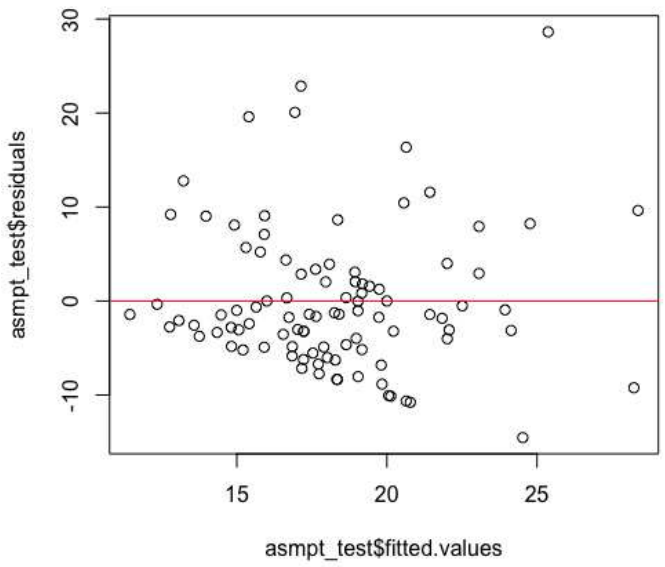
	Coefficient	BootMean	BootSE	BootLLCI	BootULCI
Using Generalised Narcissism as the Moderator					
Constant	2.60	2.61	0.47	1.65	3.50
Empathy	-0.78	-0.78	0.35	-1.43	-0.05
Narcissism	0.28	0.30	0.13	0.07	0.57
Age	-0.00	-0.00	0.01	-0.03	0.02
Gender	-0.40	-0.41	0.21	-0.82	0.00
Interaction	0.06	-0.13	0.50	-1.35	0.57
Using Vulnerable Narcissism as the Moderator					
Constant	2.38	2.38	0.47	1.40	3.26
Empathy	-0.84	-0.83	0.36	-1.52	-0.11
Vulnerable Narcissism	0.27	0.29	0.12	0.07	0.56
Age	-0.00	-0.00	0.01	-0.03	0.03
Gender	-0.30	-0.31	0.21	-0.71	0.09
Interaction	0.16	-0.02	0.49	-1.19	0.73
Using Grandiose Narcissism as the Moderator					
Constant	2.67	2.66	0.47	1.70	3.56
Empathy	-0.77	-0.73	0.35	-1.41	-0.01
Grandiose Narcissism	0.21	0.23	0.11	0.02	0.45
Age	-0.01	-0.00	0.01	-0.03	0.02
Gender	-0.43	-0.43	0.22	-0.86	-0.00
Interaction	-0.01	-0.14	0.41	-1.14	0.50

Note. Generalised Narcissism: R = 0.45, F(5, 88) = 4.56, p = .001, explaining 20.59% (R-squared) of the variance with an MSE of 57.32. Vulnerable Narcissism: R = .45, F(5, 88) = 4.53, p = .001, explaining 20.47% (R-squared) of the variance with an MSE of 57.41. Grandiose Narcissism: R = .44, F(5, 88) = 4.21, p = .002, explaining 19.29% (R-squared) of the variance with an MSE of 58.26.

Appendix D

Testing Assumptions of the Moderation Model

Linearity



Independence

Durbin-Watson Test:

- DW = 1.89, p-value = .290
- The p-value is not significant, indicating no evidence of autocorrelation in the residuals. This suggests that the assumption of independence is met.

Homoscedasticity

Breusch-Pagan Test:

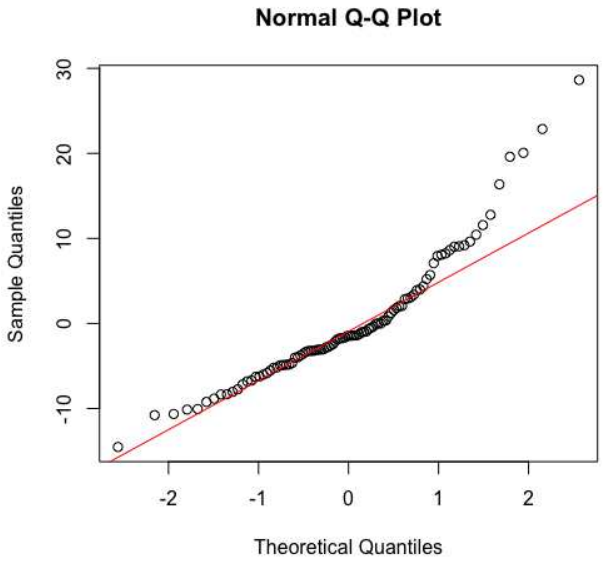
- BP = 4.12, df = 3, p-value = .249
- The p-value is not significant, suggesting no evidence of heteroscedasticity.

Non-constant Variance Test (ncvTest):

- Chi-square = 8.02, Df = 1, p = .005
- The p-value is significant, indicating evidence of heteroscedasticity.

Normality of Residuals

Q-Q Plot:



- Visual inspection of the Q-Q plot is necessary, but if it deviates significantly from the line, it indicates non-normality.

Shapiro-Wilk Test:

- W = 0.91022, p-value = 6.548e-06
- The p-value is highly significant, indicating that the residuals are not normally distributed.

Kolmogorov-Smirnov Test:

- D = 0.14753, p-value = 0.02745
- The p-value is significant, further confirming that the residuals are not normally distributed.

Multicollinearity

Variance Inflation Factor (VIF):

- GVIFs are 1 for both predictors, indicating no evidence of multicollinearity.

Appendix E

Simple Slopes Plot of the Moderation Analysis With Sensitivity Measures