Dark Tetrad of Personality and Lying in a Cheating Task

Name: Uveys Uysal
Student number: 420501
Email: uveys-uysal@hotmail.com
Master: Clinical Psychology
Supervisor: Dr. Eric Rassin
Second supervisor: Dr. Irena Boskovic
EC : 32
Word count: 6911
Instituut voor psychologie, Faculteit der Sociale Wetenschappen
Erasmus University Rotterdam

How much do people lie, who lies, and who lies a lot? Deception literature often seeks to answer these questions. Dark personality traits such as Machiavellianism, narcissism, psychopathy, and sadism have been shown to play a role in who lies and who lies a lot. We investigated which dark personality traits predicted lying and lying frequency in a cheating task. Participants completed a survey with the short dark tetrad (SD4) and a numerical matrix task. Participants had the option to lie about their performance in the matrix task while they could not be caught. We found Machiavellianism to predict lying while psychopathy predicted lying frequency when we controlled for the effect of age and gender. Sadism did also predict lying but only when not controlling for age. The role of dark personalities in cheating task lying was different with the use of the dark tetrad compared to studies which used the dark triad.

Lying is a phenomenon everyone may have experienced, whether it was by telling a lie or getting lied to. Lying can be defined as a falsely believed statement made to a receiver with the intention that the statement is believed to be true by the receiver (Mahon, 2008). It can be experienced daily while the regular occurrence of lying can be explained by viewing lying as a part of social interaction and having value in human survival (Ford et al., 1988). Studies within deception literature often research how regular the occurrence of lying is. In a pen-and-paper diary-based study by DePaulo et al. (1996), participants reported their average lies per day. For college students this was 1.96 lies per day (*SD* = 1.63, *N* = 77) whilst for community members this was .97 lies per day (SD = .98, *N* = 70). Participants in this study did not regard their lies as serious nor planned them out or worried about being caught, which complies with the notion of lying being part of social interaction (DePaulo, 1996). This study was replicated by Hancock et al. (2004) where students in this study reported an average lies per day of 1.58 (*SD* = 1.02, *N* = 28). In George and Robb (2008) the study was replicated by reporting on a personal digital assistant rather than pen and paper. Students reported an average lies per day rate of .59 (*SD* = .37, *N* = 25). Consequently, the results and the replications of DePaulo et al. (1996) have become a popular reference for lying frequency in deception literature. Ultimately, contributing to the notion that 'Everybody Lies' (DePaulo, 2004).

More recent studies however show that the notion of 'Everybody lies' is not nuanced enough when describing the regular occurrence of lying. In a mass survey by Serota et al. (2010) about lies told in the last 24 hours, participants reported an average of 1.65 lies per day (*SD* = 4.45, *N* = 998). Although the mean of the lying frequency was comparable to DePaulo et al. (1996), the positively skewed distribution of the responses led to a broader conclusion. The majority of lies were told by a small portion of the sample, a few prolific liars. Therefore this study concluded that an average lying frequency rate can be rather misleading. Additionally, approximately 60% of participants in Serota et al. (2010) reported no lies at all. Further replications of this study found a comparable positively skewed distribution of lying frequency contributing to the notion of 'prolific liars' (Halevy et al., 2014; Serota & Levine, 2015)

In Smith et al. (2014), these two notions or rather perspectives were classified and compared by analyzing lying frequency in text messages. Firstly, in the everyday lies perspective, most people are expected to tell only a small number of lies (e.g. DePaulo et al., 1996; Hancock et al., 2004). According to this perspective, the distribution of lying frequency is relatively normal with most people telling at least one lie. Secondly, in the prolific liars perspective, a few people are expected to tell the majority of the lies (Serota et al., 2010). Contrarywise according to this perspective, the distribution of lying frequency is positively skewed with most people telling no lie.

Smith et al. (2014) found that the majority of their participants (77%) reported at least one lie, supporting the everyday lie perspective. Nevertheless, the distribution of the lying frequency was positively skewed with 4% of the sample labeled as prolific liars accounting for 15% of the total lies in the sample. Therefore Smith et al. (2014) reported that a majority of individuals lie in some moderate amount, whereas there is a small number of prolific liars who lie frequently. Eventually concluding that the pattern of lying frequency lies between the everyday lies perspective and the prolific liar's perspective. The present study will research whether a current pattern of lying frequency is in line with the everyday lie perspective, the prolific liars perspective, or between the two.

Studies in deception literature have also compared types of lies, the content of lies, and recipients of lies to get a better understanding of the lying frequency differences reported by studies. Specifically, Smith et al (2014) reported more lying for smaller lies and when people talked about their plans, actions, and whereabouts. In comparison, fewer lies were reported when people talked about their achievements and when people were in a closer relationship with the recipient. Studies have also shown more lying when people are benefitting others (Gino et al., 2013) or when there is a charitable gain (Lewis et al., 2012). Furthermore, situational factors also seem to make a difference in lying frequency (Ariely, 2015). Studies have shown that more lying occurs in the morning compared to the afternoon (Kouchaki & Smith, 2014). More lying is also observable when people are under time pressure (Shalvi et al., 2012) or when self-control resources are depleted (Gino et al., 2011).

Additionally, with the ever-growing popularity of digital communication methods, research into lying frequency within these methods has also been popularized. Hancock et al. (2004) reported most average lies per day for social interactions by phone and least for social interactions by email. A common form of lies in text messaging is *butler lies*, in which lies are about avoiding or delaying social interaction (Hancock et al., 2009). For this reason, Smith et al. (2014) argue that the prevalence of butler lies in text messages could be the reason for the majority of their participants telling at least one lie. This would provide support for deception being more frequent in computer-mediated contexts compared to face-to-face contexts (Drouin et al., 2016; Toma et al., 2016).

A rightful concern in deception literature is the validity of self-report instruments to measure deception or lying. To put it simply, why would a frequent liar be honest about the frequency of their lies? A common solution to this concern is to operationalize lying in the form of cheating with some form of cheating task. While cheating tasks create an opportunity to lie, the chance of getting caught is constantly zero (Halevy et al., 2014). The incentive to lie in cheating tasks is created by offering the participants a prize for better performance. This study will therefore measure lying frequency through a cheating task.

It is notable that with the everyday lie perspective on lying frequency, situational factors and communication methods become particularly interesting. That is because if everybody lies sometimes, research should focus on which situational factors promote lying (Thaler & Sunstein, 2008). Since if everybody lies, individual differences should not have much influence in identifying lying behavior (Serota et al., 2010).

However, with studies showing support for the prolific liar's perspective, research into individual differences does become interesting. Moreover, research suggests that individual differences play a role in lying frequency but are often neglected in deception literature (Serota et al., 2010; Halevy et al., 2014). For example, a meta-analysis by Bond Jr & DePaulo (2008) suggests substantial individual differences in the ability to credibly lie of liars influence the detectability of those lies.

A popular concept used in deception literature that relates to individual differences is *the dark triad of personality*. The dark triad consists of three socially aversive personality traits: Machiavellianism, psychopathy, and narcissism (Roeser et al., 2016). Machiavellianism can be described as manipulativeness and emotional coldness (Christie & Geis, 1970). Psychopathic traits are characterized by callous affect, low empathy, and impulsivity (Hare, 1985) while narcissism can be described by a sense of superiority, grandiosity, and entitlement (Raskin & Hall, 1979). Studies have indeed shown that these dark triad traits are associated with lying (Azizli et al., 2016; Jonason et al., 2014). However, research shows that Machiavellianism, psychopathy, and narcissism relate differently to lying and lying frequency. Machiavellianism seems to be a significant predictor of lying with high stakes while narcissism and psychopathy are not (Azizli et al., 2016). Machiavellianism and psychopathy seem to be related to a greater probability to lie while narcissism is not (Baughman et al., 2014). In terms of lying frequency, individuals which score high on Machiavellianism and psychopathy measures seem to lie more (Jonason et al., 2014) or cheat more frequently in cheating tasks (Roeser et al., 2016). Also, Psychopathy specifically as a personality trait implied by psychopathic tendencies is associated with lying frequency (Halevy et al., 2014). Narcissism on the other hand does not make a difference in the number of lies (Baughman et al. 2014; Jonason et al., 2014; Roeser et al., 2016).

An even further distinction can be seen when examining the type of cheating task that is used in measuring lying and lying frequency. In Roeser et al. (2016) only psychopathy predicted lying frequency in a cheating task where lying could be done more impulsively. When in fact only Machiavellianism predicted lying in a task where lying was done more thoughtfully and while considering a victim. Additionally, in Jonason et al. (2014), psychopathy was associated with telling lies for no reason and Machiavellianism was associated with telling white lies. Therefore, these findings provided support for the impulsive nature of psychopathy in lying.

Recently the dark triad of personality was extended to also include sadism, resulting in a broader conceptualization called *the dark tetrad of personalities* (Paulhus et al., 2020). Paulhus et al. (2020) argue sadism offers to include “the intrinsic pleasure in hurting others”, a new characterization that was not covered by the original dark triad. Additionally, sadism seems to be compatible with the original dark triad since it shares callousness as one of its characterizations. Compared to psychopathy, sadism is most comparable within physical characterizations, and most distinguishable within vicarious characterization (Paulhus et al., 2020).

A popular tool throughout deception literature to assess dark triad personalities is the *short dark triad* (SD3) (Jones & Paulhus, 2014). Together with the dark tetrad, Paulhus et al. (2020) also presented the *short dark tetrad* (SD4), a successor to the SD3. When comparing these tools, the obvious difference between the SD4 and the SD3 is the addition of the sadism scale within the SD4. Still, additional changes and the effect of the sadism scale have also led to substantial changes within the other three dark personality scales (Paulhus et al., 2020). Firstly, each scale in the SD4 consists of 7 items, whereas in the case of the SD3 each scale consisted of 9 items. Secondly, the SD4 unlike the SD3 does not contain any reversed items. Lastly, psychopathy and Machiavellianism are more distinguishable in the SD4 compared to the SD3 (Paulhus et al., 2020). The mean intercorrelation between psychopathy and Machiavellianism in the SD3 was .47 (Jones & Paulhus, 2014), whereas it is .31 in the SD4 (Paulhus et al., 2020). Paulhus et al. (2020) also address individual participant differences and advise possibly controlling for them when administering the SD4. This is because age is negatively related to the dark tetrad of personality (Paulhus et al., 2020) and males overall score higher on dark traits compared to females (Forsyth et al., 2021; Paulhus et al., 2020).

Altogether, some dark triad personality traits have been shown to have a place in deception literature by illustrating the ability to predict lying and lying frequency. This applies to Machiavellianism and psychopathy and not to narcissism. Most studies showed this with self-report measures for lying and lying frequency, while only some used cheating tasks. With the relatively new emergence of the dark tetrad of personality, the roles of Machiavellianism and psychopathy in lying and lying frequency have not yet been replicated with the SD4. Additionally, the role of sadism in lying and lying frequency has not been researched in a dark tetrad manner. There have been studies that measured lying frequency while also including sadism in a dark tetrad manner such as Forsyth, et al. (2021). However dark personalities were assessed by using the SD3 in conjunction with a standalone sadism measure.

Therefore, the present study researched besides the pattern of lying frequency, whether lying and lying frequency could be predicted with the dark tetrad of personalities using the SD4. An operational cheating task was used to measure lying and lying frequency where lying was manipulated to be done impulsively. We also included age and gender as potential covariates after replicating their associations with the dark tetrad. Results for Machiavellianism, psychopathy, and narcissism were compared with findings from studies that utilized the SD3 to predict lying or lying frequency. Furthermore, we researched if the role of sadism in lying and lying frequency was comparable to the roles of Machiavellianism and psychopathy from dark tetrad studies.

With these points in mind, this study aimed to provide insight into the relationship between the dark tetrad of personality and lying in a subclinical sample. The results will eventually contribute to the understanding of which individual differences predict whether people lie or not and which individual differences predict how much liars lie. Our study will also add to the body of literature that used cheating tasks to measure lying.

Method

Participants. Data was collected with the Erasmus Matrix Project (EMP), an online survey made with Qualtrics (2005). In total 961 survey entries from participants were registered. After removing incomplete entries, 715 participants remained. Followingly, only 550 participants successfully answered on the cheating task. Therefore, the study included N = 715 Dutch and English speaking participants but for results considering the cheating task a sample size of n = 550 was used. Participants were requested but not required to report their age and gender. This resulted in 23 participants not reporting their age while 6 of them neither reported their gender. Data from these participants were preserved but removed when analyses were run that considered age and gender respectively. Analyses that considered gender were only run on data from participants who identified as male or female. Mean age was 30.80 years (SD = 13.02). The sample mostly consisted of volunteers recruited by interns which were supervised by Dr. Rassin. Additionally, Student participants were recruited from the Erasmus University Rotterdam. Table 1 shows the subsample sizes for gender, user language, and recruitment. After full completion students were rewarded with 15 minutes of subject credits. All participants had a chance to win a €10,- gift card. In total 10 gift cards were rewarded.

Table 1

Participant Demographics

	Sample	
Characteristic	*n*	%
Gender		
Male	238	33.3
Female	462	64.6
Non-binary / third gender	5	0.7
Prefer not to say	4	0.6
User Language		
Dutch	478	66.9
English	237	33.2
Recruitment		
Erasmus University Students	80	11.2
Voluntary	635	88.8
Total	715	

Measures. *Short Dark Triad 4 (SD4; Paulhus et al., 2020).* The SD4 is a 28-item measure to assess dark tetrad personality traits. Within the 28-items, 1-7 measure Machiavellianism, 8-14 measure narcissism, 15-21 measure psychopathy, and 22-28 measure sadism. Each item is rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The Dutch translation of the SD4 was provided by Van Dongen et al. (2022). The SD4 demonstrates high internal consistency for Machiavellianism (α = .78), narcissism (α = .83), psychopathy (α = .82), and sadism (α = .82; *Paulhus et al., 2020*). Scores on the SD4 were averaged per subscale.

Matrix task. The matrix task, as originally introduced by Mazar et al. (2008) is a way to measure dishonest behavior utilizing a cheating task. Participants in this study were presented with a total of 20 matrices. Each matrix consisted of 12 random numbers divided into four rows and three columns (see figure 1). Numbers had two decimals or one when the last decimal was zero.

Figure 1

Example of a matrix used in the matrix task to measure dishonest behavior

3.12	9.34	1.99
9.45	5.21	9.35
8.01	8.72	1.96
2.66	8.25	9.88

Note. The correct combination is 8.01 and 1.99.

For each matrix, participants were asked to find the combination of two numbers that added up to 10. While searching for this combination a timer was running to limit time spent on each matrix to 30 seconds. Additionally, participants were requested to keep count of the total number of matrices they solved. It was advised to keep a score with pen and paper. At the end of the matrix task, participants were asked to input the total number of matrices they had managed to solve. However, participants were not required to provide the numerical combinations themselves. Consequently, this design takes away the option to check the answers and the chance of getting caught is zero. Only the first 13 out of 20 matrices were solvable. Which meant participants who reported to have solved more than 13 matrices were cheating. The number of solved matrices above 13 was therefore used to measure lying and a lying frequency. Participants did have an incentive to cheat because they were told that the top 10 best performers on the matrix tasks would have a higher chance to win a gift card.

An overview of all the matrices that were used in this study can be found in Appendix A. All of these matrices were randomly generated by a custom R script. A copy of this script can be found in Appendix B.

Procedure. Participants were presented with two parts in the EMP survey. Part 1 contained questions about consent, age, gender, and a questionnaire with 28 SD4 items. Part 2 contained the matrix task with 20 matrices. Only part 2 was time limited. All items, tasks, as well as instructions, were available in Dutch and English. Participants had the option to access the questionnaire online from their own desired location and device. All participants gave informed consent. However, participants were not informed on the exact nature of the measurement tools nor an the incentive to lie. They were told the study was about solving numeric puzzles and the association with personality traits. If participants desired more information about the nature of the study and measures, they could mail the thesis supervisor.

Analysis. Associations between age and SD4 subscale means were examined with Spearman's correlation coefficients. Differences between males and females on average SD4 subscales were assessed with independent *t*-tests and Welch *t*-tests.

Binomial logistic regression analyses were used to examine predictors of the likelihood to lie. Firstly, four models were run with each SD4 subscale respectively. Followingly, age and gender were included in all four sperate models as covariates. Next, all SD4 subscales were included in one model. Finally, alongside the SD4 subscales, age and gender were included as covariates in one model.

Poisson regression analyses were used to examine predictors of lying frequency within the matrix task. Here two models were run on data from participants that lied. One model included all SD4 subscales as predictors, while the second model included the SD4 subscales, age and gender.

Analyses were conducted using R (2022). Appendix C contains a R script with all steps that were taken during data analysis. Statistical significance was assessed using $\alpha = 0.05$.

Results

Participants' age was negatively correlated with scores on Machiavellianism, $r_s = -.15$, $p < .001$, and sadism, $r_s = -.23$, $p < .001$. Participants' age was not correlated with narcissism or psychopathy. Spearman's rank correlation was used since the assumption of linearity was violated between participants' age and the SD4 subscales while also participants' age was non normally distributed.

Males scored on average significantly higher on all SD4 subscales compared to females (Table 2). As assessed by the Levene's test for equality of variances, the assumption of homogeneity of variances was violated for psychopathy scores between males and females, $F(1, 698) = 23.05$, $p <.001$, and sadism scores between males and females, $F(1, 698) = 15.96$, $p <.001$. Therefore, Welch's t-test was used for the psychopathy and sadism subscales.

Table 2

Independent t-test results for average scores on SD4 subscales between males and females

	Males (n =192)	Females (n = 351)			
	M (*SD*)	*M* (*SD*)	*t*	*df*	*p*
Machiavellianism	3.24 (0.54)	3.02 (0.51)	5.29	698	< .001
Narcissism	2.96 (0.65)	2.75 (0.63)	4.19	698	< .001
Psychopathy	2.30 (0.70)	1.94 (0.55)	6.92	390.22	< .001
Sadism	2.74 (0.76)	2.06 (0.61)	12.01	398.23	< .001

Note. Welch's t-test for psychopathy and sadism.

In total, 52 out of 550 participants reported having solved more than 13 matrices. This is indicated by the "impossible" answers in figure 2. That means that 9.5% of participants were confirmed to have lied on the matrix task. The average number of lies for the participants who did lie was 3.69 (*SD* = 2.28). There was no significant difference in the average number of lies between males who lied and females who lied. Average scores on the SD4 subscales for participants who lied and who did not lie can be seen in table 3.

Figure 2

Answers from participants on the matrix task

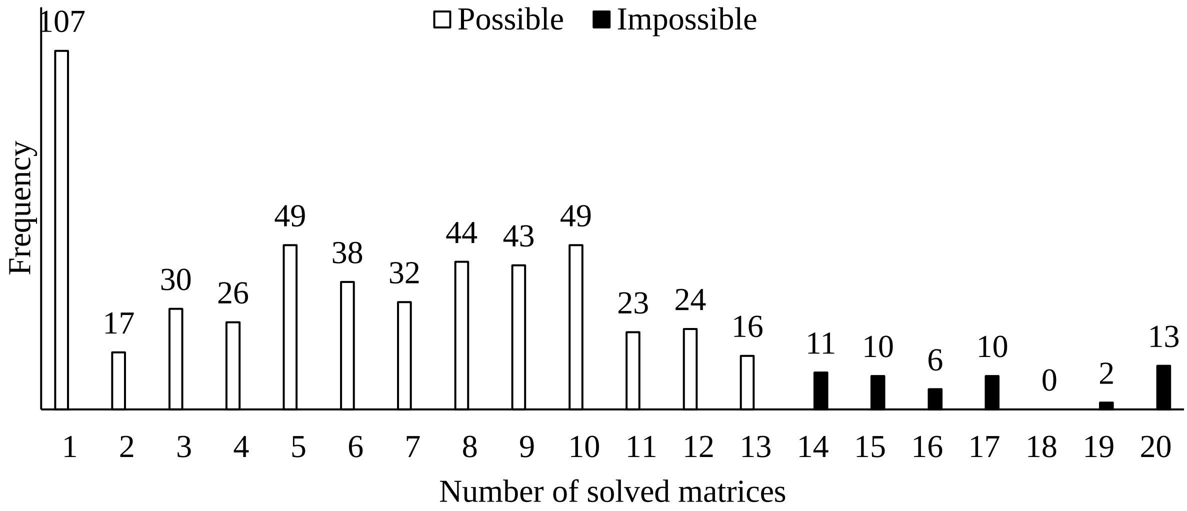

Table 3

Average scores on SD4 subscales between participants who lied and who did not lie

	No Lie (n = 498)	Lie (n = 52)
	M (*SD*)	*M* (*SD*)
Machiavellianism	3.07 (0.52)	3.39 (0.43)
Narcissism	2.82 (0.64)	3.01 (0.69)
Psychopathy	2.05 (0.63)	2.37 (0.77)
Sadism	2.28 (0.72)	2.76 (0.83)

Lying. Answers from the matrix task were recoded as “No Lie” (1-13) or “Lie” (14-20). Logistic regressions were performed to examine the effect of Machiavellianism, narcissism, psychopathy, and sadism on the likelihood that a participant was lying. Linearity of the SD4 subscales with respect to the logit of the likelihood to lie was assessed with partial regression plots and the Box-Tidwell (1962) method. Based on these assessments, all SD4 subscales were found to be linearly related to the logit of the likelihood to lie. There was no evidence of multicollinearity, as assessed by variance inflation factor (VIF) values less than 10. Outlier diagnostics indicated some influential scores. All scores were however included in the analyses because scores were compliant with the ranges of our measuring instruments.

Four models were run for each SD4 subscale separately. Machiavellianism, narcissism, psychopathy, and sadism were each found to be significant predictors of the likelihood to lie within their respective models (Table 4).

Table 4

Logistic Regressions Predicting Likelihood to lie based on SD4 subscales separately

	B	SE	Wald Z	*p*	$\chi 2$	*p*	Nagelkerke R^2
Machiavellianism	1.21	0.29	4.13	<.001	17.94	<.001	0.07
Narcissism	0.46	0.22	2.05	.040	4.23	.040	0.02
Psychopathy	0.70	0.21	3.35	<.001	10.94	<.001	0.04
Sadism	0.81	0.19	4.27	<.001	18.36	<.001	0.07

Note. Coefficients are from separate models. Each model has one degree of freedom.

Afterwards, logistic regression models for the effect of SD4 subscales separately on the likelihood to lie when controlling for the effect of age and gender were found to be statically significant (Table 5).

Table 5

Model statistics for Logistic Regressions Predicting Likelihood to lie based on SD4 subscales separately including Age and Gender

	$\chi 2$	p	Nagelkerke R^2
Machiavellianism	24.90	<.001	0.10
Narcissism	16.87	<.001	0.07
Psychopathy	19.43	<.001	0.08
Sadism	20.48	<.001	0.08

Note. Subscale names refer to models not coefficients. Each model has three degrees of freedom.

From these four separate models only narcissism was not a significant predictor of the likelihood to lie when controlling for the effect of age and gender (Table 6). Machiavellianism, psychopathy, and sadism all were significant predictors of the likelihood to lie when controlling for the effect of age and gender.

Table 6

Logistic Regressions Predicting Likelihood to lie based on SD4 subscales separately including Age and Gender

	B	SE	Wald Z	*p*
Machiavellianism				
Machiavellianism	0.99	0.31	3.16	.002*
Age	-0.04	0.02	-2.03	.043*
Gender	-0.65	0.326	-2.03	.042*
Narcissism				
Narcissism	0.35	0.23	1.53	.127
Age	-0.05	0.02	-2.50	.012*
Gender	-0.76	0.31	-2.41	.016*
Psychopathy				
Psychopathy	0.50	0.21	2.22	.026*
Age	-0.04	0.02	-2.34	.019*
Gender	-0.62	0.33	-1.91	.057
Sadism				
Sadism	0.57	0.23	2.42	.015*
Age	-0.03	0.02	-1.74	.082
Gender	-0.40	0.36	-1.10	.271

Note. Gender is for females compared to males. $^{*}p < .05$.

Next, a logistic regression model with every SD4 subscale analyzed jointly was found to be statistically significant, χ2(4) = 26.31, p = <.001, Nagelkerke R^2 = 0.11. When controlling for shared explained variance between the SD4 subscales, only Machiavellianism and sadism were found to be significant predictors of the likelihood to lie (Table 7). An increase in Machiavellianism and an increase in sadism were associated with an increase in the likelihood to lie. Narcissism and psychopathy were not significant predictors of the likelihood to lie.

Table 7

Logistic Regression Predicting Likelihood to lie based on SD4 subscales, Age, and Gender

	Without covariates				With covariates			
	B	SE	Wald Z	*p*	*B*	SE	Wald Z	*p*
Machiavellianism	0.93	0.32	2.89	.004*	0.85	0.32	2.67	.008*
Narcissism	-0.04	0.26	-0.14	.889	0.00	0.26	0.01	.945
Psychopathy	0.24	0.27	0.88	.378	0.26	0.27	0.97	.333
Sadism	0.52	0.24	2.19	.028*	0.32	0.27	1.20	.231
Age	-	-	-	-	-0.03	0.02	-1.58	.114
Gender	-	-	-	-	-0.32	0.36	-0.88	.381

Note. Gender is for females compared to males. $^*p < .05$.

Lastly, a logistic regression model for the effect of every SD4 subscale on the likelihood to lie when controlling for the effect of age and gender was found to be statically significant, χ2(6) = 29.51, p = <.001, Nagelkerke R^2 = 0.12. Only Machiavellianism was a significant predictor of the likelihood to lie when we controlled for the effect of age and gender (Table 7). However, we found the model with covariates to not significantly predict more compared to the model without covariates, χ2(2) = 3.20, p = .201. Psychopathy, narcissism, and sadism were not significant predictors of the likelihood to lie when controlling for the effect of age and gender.

Lying frequency. Poisson regression analyses were run to predict lying frequency based on the SD4 subscales. Dispersion within our model was assessed by tests for overdispersion as described by Cameron and Trivedi (1990). Although some overdispersion was observed in our model, equidispersion was assumed, ϕ = 1.25, Z = 1.24, p = .107. Only psychopathy was found to be a significant predictor of lying frequency (Table 8). Higher psychopathy scores resulted in a higher number of lies. Age and gender did not change the association between psychopathy and the frequency of lies. The model with covariates did also not significantly predict more compared to the model without covariates, $\chi 2(2) = -0.53$, p = .767. Machiavellianism, narcissism, and sadism did not significantly predict lying frequency.

Table 8

Poisson Regression Predicting lying frequency on SD4 subscales, Age, and Gender

	Without covariates				With covariates			
	B	SE	Wald Z	*p*	*B*	SE	Wald Z	*p*
Machiavellianism	0.11	0.19	0.58	.561	0.10	0.19	0.51	.612
Narcissism	0.17	0.14	1.17	.243	0.17	0.14	1.21	.226
Psychopathy	0.26	0.13	2.02	.043*	0.28	0.13	2.13	.033*
Sadism	-0.15	0.12	-1.32	.188	-0.17	0.13	-1.28	.200
Age	-	-	-	-	-0.01	0.01	-0.66	.507
Gender	-	-	-	-	0.05	0.17	0.29	.771

Note. Gender is for females compared to males. $^{*}p < .05$.

Discussion

The present study started with researching the pattern of lying frequency. Since we only confirmed lies with 9.5% of participants, our findings would be most in line with the prolific liar's perspective. Compared to the previous studies on lying frequency, our sample could be characterized as rather honest. It should however be noted that, in contrast to the previously mentioned studies on lying frequency, we did not use self-reporting to measure lying. The use of cheating tasks instead of self-reports could be a reason why we did measure a low number of lies. Indeed, Roeser et al. (2016) also experienced a low frequency of lies within their sample when they administered a matrix task.

Another possible explanation for the low frequency of lies is the way in which participants were recruited. Most of the participants were volunteers that were asked to complete the survey by interns of the thesis supervisor. It is possible participants were less inclined to cheat on the matrix task when they were doing the recruiter a favor.

It is also possible that our sample was actually not honest. This is because of a possible flaw within our design of the matrix task in which liars are confirmed. As we mentioned, we only confirmed lying when participants reported having solved more than 13 matrices. Since participants did not indicate which matrix they had solved, participants could have indicated they had solved impossible matrices while they reported less than 13 solved matrices. This would mean that the actual number of lies is higher than we measured. This uncertainty could have been possibly remedied by asking participants whether they had solved a matrix after each matrix. However, this was not done to prevent overcontrolling which in turn could have decreased the incentive to lie. Roeser et al. (2016) did confirm lies after every matrix but still reported having an extremely honest sample. Unfortunately, they did not report a percentage of liars for their matrix task.

Additionally, a meta-analysis by Gerlach et al. (2019) compared the results of 101 matrix task studies to studies which used different cheating tasks. This meta-analysis concluded that although the matrix tasks had a comparable rate of liars (48%) to other cheating tasks, participants in the matrix task cheated to a lower degree. Gerlach et al. (2019) discussed that this was possibly because top performances in a matrix task are less likely to be obtained by honest behaviour compared to other cheating tasks.

Our study replicated the negative association between age and Machiavellianism as well as sadism. However, we did not replicate any association between age and psychopathy. Neither did we for age and sadism. This inconsistency in findings can be explained by the reported ages in our sample. Over 70% of our sample was younger than 30 whilst we collected data from everyone older than 18. This meant we had relatively fewer data from elderly participants. Our results did replicate the findings of Paulhus et al. (2020) that males scored on average higher on all dark tetrad personalities compared to females. This already was the case with the dark triad of personality (Azizli et al., 2015; Jonason et al., 2014), which confirms that sex differences in dark personalities is not different in the dark tetrad.

Followingly, we sought to answer whether lying could be predicted with the dark tetrad of personalities. Machiavellianism and sadism predicted lying in the matrix task, where higher subscale scores resulted in higher chances to lie. The role of Machiavellianism in predicting lying was already fairly established in the literature (Baughman et al., 2014; Roeser et al., 2016). The absence of psychopathy and the presence of sadism, however, is not entirely in line with previous studies. In Roeser et al. (2016) it was also shown that out of all the dark triad personalities, only Machiavellianism predicted lying in a cheating task. This cheating task however required high cognitive effort in contrast to a matrix task where it is done simply and impulsively (Roeser et al., 2016). Based on our implementation of the matrix task it could have been that participants did not lie as impulsively as we originally intended. After they were introduced to the matrix mask, participants in our study had a total of ten minutes before they had to report their solved matrices. It could have been that some of them already planned to lie during this time.

Another possible explanation for the absence of psychopathy is the presence of sadism itself. Previous studies have shown that Machiavellianism and psychopathy are more similar in their cheating strategy than narcissism (Jonason et al., 2014). With the inclusion of sadism in the dark tetrad, Machiavellianism and psychopathy are more distinguishable compared to the dark triad (Paulhus et al., 2020). This could mean that the similarity in cheating between Machiavellianism and psychopathy in the dark tetrad is not as apparent due to the inclusion of sadism.

Additionally, Paulhus et al. (2020) explain that in contrast to angry, fearless, and impulsive behavior in psychopathy, sadism is characterized by a preference for vicarious rewards where a safe distance is measured. Lying within our study is better described with the latter. It is therefore possible that the absence of this distinction in the dark triad resulted in the effect of psychopathy for lying, while the presence of this distinction resulted in the effect of sadism for lying.

Sadism did not predict lying when we controlled for age and gender. We found this to be a result of the inclusion of age since sadism did predict lying when only gender was included. We believe this to be a result of the relatively high correlation between sadism and age compared to the other dark tetrad personalities and age. Including covariates, however, did not result in a significantly better prediction of lying.

Taken together, we conclude that using the dark tetrad, the role of Machiavellianism in lying is the same compared to studies which used the dark triad. Whereas the role of psychopathy can be seen as different or even replaced by sadism in the dark tetrad.

In terms of lying frequency, only psychopathy predicted the frequency of lies within the matrix task. Higher scores on psychopathy resulted in more lies. These findings are in line with the finding of Roeser et al. (2016) where compared to other dark tetrad personalities, only psychopathy predicted lying frequency in a matrix task. Yet our findings do not support previous studies which found Machiavellianism to also predict lying frequency (Azizli et al., 2015; Jonason et al. 2014). These studies however made use of self-report measures instead of cheating tasks to measure lying frequency. Additionally, as reported by previous studies narcissism did not predict lying or lying frequency in our study. This leads us to confirm that the role of narcissism within the dark tetrad for lying and lying frequency is the same compared to studies that used the dark triad.

The effect of the dark tetrad subscales, excluding sadism, on lying and lying frequency did not change when we controlled for age and gender. We also did not find an effect of Age or gender on lying or lying frequency. Gerlach et al. (2019) did confirm an effect of age and gender on dishonest behaviour in cheating tasks, where men behaved slightly more dishonest and older participants were associated with less dishonest behaviour. However, compared to other cheating tasks, the effect of gender was smallest in the matrix task while the effect of age was not present in the matrix task (Gerlach et al., 2019). This may explain the absences of these effects in our study.

Comparisons between our results and studies that utilized self-reporting measures should be done carefully. Importantly we want to reiterate that these findings on lying and lying frequency are based on a relatively small sample of liars. Our conclusions on lying and lying frequency were not based on the separate dark tetrad subscale analyses since we were only concerned with unique explained variance.

Concludingly, the present study provides support for the prolific liars perspective and the role of Machiavellianism in lying when controlling for age and gender. We also provide support for sadism in lying without controlling for age and the role of psychopathy in lying frequency when controlling for age and gender. The role of dark personalities in cheating task lying was different with the use of the dark tetrad compared to studies which used the dark triad. Ultimately, this adds to the better understanding of which dark personalities play a role in lying and lying frequency. Future research with the use of cheating tasks could shed light on a more nuanced understanding of which dark personalities are associated with various types of lies.

References

Ariely, D. (2015). The Honest Truth about Dishonesty: How We Lie to Everyone--Especially Ourselves.

Azizli, N., Atkinson, B. E., Baughman, H. M., Chin, K., Vernon, P. A., Harris, E., & Veselka, L. (2016). Lies and crimes: Dark Triad, misconduct, and high-stakes deception. *Personality and Individual Differences*, *89*, 34-39.

Baughman, H. M., Jonason, P. K., Lyons, M., & Vernon, P. A. (2014). Liar liar pants on fire: Cheater strategies linked to the Dark Triad. *Personality and Individual Differences*, *71*, 35-38.

Bond Jr, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*(4), 477.

Box, G. E., & Tidwell, P. W. (1962). Transformation of the independent variables. Technometrics, 4(4), 531-550.

Cameron, A. C., & Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model. *Journal of Econometrics*, *46*(3), 347-364.

Christie, R., & Geis, F. L. (1970). Studies in Machiavellianism New York: Academic Press.

DePaulo, B. M. (2004). The many faces of lies. *The Social Psychology of Good and Evil*, 303-326.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*, 979.

Drouin, M., Miller, D., Wehle, S. M., & Hernandez, E. (2016). Why do people lie online? "Because everyone lies on the internet". *Computers in Human Behavior*, *64*, 134-142.

Ford, C. V., King, B. H., & Hollender, M. H. (1988). Lies and liars: psychiatric aspects of prevarication. *The American Journal of Psychiatry*.

Forsyth, L., Anglim, J., March, E., & Bilobrk, B. (2021). Dark Tetrad personality traits and the propensity to lie across multiple contexts. *Personality and Individual Differences*, *177*, 110792.

George, J. F., & Robb, A. (2008). Deception and computer-mediated communication in daily life. *Communication Reports*, *21*, 92-103.

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, *145*, 1-44.

Gino, F., Ayal, S., & Ariely, D. (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of Economic Behavior & Organization*, *93*, 285-292.

Gino, F., Schweitzer, M. E., Mead, N. L., & Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, *115*, 191-203.

Halevy, R., Shalvi, S., & Verschuere, B. (2014). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research*, *40*, 54-72.

Hancock, J., Birnholtz, J., Bazarova, N., Guillory, J., Perlin, J., & Amos, B. (2009, April). Butler lies: awareness, deception and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 517-526).

Hancock, J. T., Thom-Santelli, J., & Ritchie, T. (2004, April). Deception and design: The impact of communication technology on lying behavior. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 129-134).

Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, *53*, 7.

Hart, C. L., Lemon, R., Curtis, D. A., & Griffith, J. D. (2020). Personality traits associated with various forms of lying. *Psychological Studies*, *65*, 239-246.

Heyman, T., Vankrunkelsven, H., Voorspoels, W., White, A., Storms, G., Verheyen, S., ... & Inbar, Y. (2020). When cheating is an honest mistake: A critical evaluation of the matrix task as a measure of dishonesty. *Collabra: Psychology*, *6*.

Jakobwitz, S., & Egan, V. (2006). The dark triad and normal personality traits. *Personality and Individual differences*, *40*, 331-339.

Jonason, P. K., Lyons, M., Baughman, H. M., & Vernon, P. A. (2014). What a tangled web we weave: The Dark Triad traits and deception. *Personality and Individual Differences*, *70*, 117-119.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short dark triad (SD3) a brief measure of dark personality traits. *Assessment*, *21*, 28-41.

Kouchaki, M., & Smith, I. H. (2014). The morning morality effect: The influence of time of day on unethical behavior. *Psychological Science*, *25*, 95-102.

Lewis, A., Bardis, A., Flint, C., Mason, C., Smith, N., Tickle, C., & Zinser, J. (2012). Drawing the line somewhere: An experimental study of moral compromise. *Journal of Economic Psychology*, *33*(4), 718-725.

Mahon, J. E. (2008). Two definitions of lying. *International Journal of Applied Philosophy*, *22*(2), 211-230.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*, 633-644.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological Assessment*, *23*, 523.

Paulhus, D. L., Buckels, E. E., Trapnell, P. D., & Jones, D. N. (2020). Screening for dark personalities. *European Journal of Psychological Assessment*.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*, 556-563.

Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological reports*.

Roeser, K., McGregor, V. E., Stegmaier, S., Mathew, J., Kübler, A., & Meule, A. (2016). The Dark Triad of personality and unethical behavior at different times of day. *Personality and Individual Differences*, *88*, 73-77.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research*, *36*, 2-25.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology*, *34*, 138-157.

Shalvi, S., Eldar, O., & Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science*, *23*, 1264-1270.

Smith, M. E., Hancock, J. T., Reynolds, L., & Birnholtz, J. (2014). Everyday deception or a few prolific liars? The prevalence of lies in text messaging. *Computers in Human Behavior*, *41*, 220-227.

Thaler, R. H., & Sunstein, C. R. (2008). Nudge: improving decisions about health. *Wealth, and Happiness*, *6*, 14-38.

Toma, C. L., Jiang, L. C., & Hancock, J. T. (2016). Lies in the eye of the beholder: Self-other asymmetry in beliefs about deception across interpersonal media. *Communication Research.* Advance online publication.

Van Dongen, J.D.M., De Roos, M.S. & Rassin, E. (2022). Nederlandse vertaling SD4. Erasmus Universiteit Rotterdam

Appendix A

3.76	7.19	9.37
0.63	9.19	2.58
0.38	4.48	6.64
9.41	6.29	4.31

3.76	7.19	9.37
0.63	9.19	2.58
0.38	4.48	6.64
9.41	6.29	4.31

1

4.65	3.63	1.54
7.14	3.13	7.64
6.33	4.65	2.36
3.91	2.18	5.45

4.65	3.63	1.54
7.14	3.13	7.64
6.33	4.65	2.36
3.91	2.18	5.45

2

0.04	8.15	1.09
3.11	4.85	6.36
8.89	1.67	1.97
8.8	8.03	7.2

0.04	8.15	1.09
3.11	4.85	6.36
8.89	1.67	1.97
8.8	8.03	7.2

3

8.42	1.94	1.34
8.06	3.06	7.61
3.84	4.91	1.99
6.93	9.29	3.51

8.42	1.94	1.34
8.06	3.06	7.61
3.84	4.91	1.99
6.93	9.29	3.51

4

2.36	6.26	3.38
7.51	3.49	6.5
7.05	7.33	5.8
6.62	8.31	8.93

2.36	6.26	3.38
7.51	3.49	6.5
7.05	7.33	5.8
6.62	8.31	8.93

5

3.96	7.7	3.49
1.41	8.62	9.31
8.73	9.16	9.27
2.25	0.69	5.89

3.96	7.7	3.49
1.41	8.62	9.31
8.73	9.16	9.27
2.25	0.69	5.89

6

7.96	7.98	9.21
2.73	4.36	5.64
8.1	8.05	0.37
5.5	1.23	3.63

7.96	7.98	9.21
2.73	4.36	5.64
8.1	8.05	0.37
5.5	1.23	3.63

7

5.51	5.63	2.07
9.66	9.23	9.04
4.94	9.71	6.67
3.33	1.4	4.18

5.51	5.63	2.07
9.66	9.23	9.04
4.94	9.71	6.67
3.33	1.4	4.18

8

7.74	8.43	3.8
5.65	4.12	7.01
6.71	8.32	2.99
6.93	8.15	9.89

7.74	8.43	3.8
5.65	4.12	7.01
6.71	8.32	2.99
6.93	8.15	9.89

9

3.95	7.2	8.57
4.52	4.49	5.35
9.69	2.43	7.54
1.71	5.51	2.53

3.95	7.2	8.57
4.52	4.49	5.35
9.69	2.43	7.54
1.71	5.51	2.53

10

6.27	5.86	3.73
4.03	7.26	5.23
4.04	1.24	1.26
9.15	6.08	3.93

6.27	5.86	3.73
4.03	7.26	5.23
4.04	1.24	1.26
9.15	6.08	3.93

11

1.73	0.31	3.72
5.62	9.8	5.72
8.27	8.43	3.16
6.79	8.47	0.06

1.73	0.31	3.72
5.62	9.8	5.72
8.27	8.43	3.16
6.79	8.47	0.06

12

2.08	1.37	7.52
9.51	6.58	1.17
0.65	4.23	8.11
6.52	9.35	8.9

2.08	1.37	7.52
9.51	6.58	1.17
0.65	4.23	8.11
6.52	9.35	8.9

13

8.56	8.98	2.57
9.47	8.79	1.43
9.84	8.73	6.73
7.71	0.96	4.77

1

6.18	3.27	7.21
2.04	2.22	7.25
2.85	8.17	1.76
2.07	3.54	9.57

2

4.14	4.2	1.21
7.01	3.27	5.48
7.16	3.95	5.97
4.65	8.7	5.82

3

3.1	8.56	6.81
0.75	6.34	9.2
4.43	7.63	2.56
5.22	6.01	0.14

4

9.43	3.94	4.52
0.15	8.78	2.66
8.3	7.13	3.91
3.56	0.6	7.58

5

0.65	9.49	3.56
7.87	5.99	7.38
9.46	0.95	5.45
9.49	6.31	9.2

6

9.63	9.91	1.46
8.63	8.83	5.88
0.42	7.4	5.85
7.52	4.54	5.11

7

Appendix B

```
# Matrix generation used in the Erasmus Matrix Project
# Created : 21 January 2022
# Author: Uveys Uysal
# E-mail: uveys-uysal@hotmail.com
# Copyright: Free to use. Please do credit U.Uysal

# matcheck = Function to check how many combinations of numbers in a vector
add to 10
matcheck <- function(mynumbers){
  # mynumbers = vector of selected digits
  # A vector of 12 random numbers between 0-10 is recommended for the
matrix task
  acount <- 0 #initial count of combinations
  for (i in 1:length(mynumbers)) {
    for (j in 1:length(mynumbers)) {
      if(i != j){
        mysum <- mynumbers[i] + mynumbers[j] #adding up 2 numbers
        if (mysum == 10) {
          cat(mynumbers[i], mynumbers[j], '\n') #print which numbers added
to 10
          acount <- acount + 1 #raise count of combinations
        }
      }
    }
  }
  return(acount)
}

######################### Solvable Matrices #########################
repeat{
  ##Loop to make a new vector of numbers until only 1 combination of
numbers adds to 10

  # create a vector of 12 random digits rounded on 2 digits
  vec <- round(runif(12, 0, 10),2)

  # create 2 answers which add to 10
  answ <- round(runif(1, 0, 10), 2)
  tansw <- 10-answ

  # create 2 random spots for the 2 answers
  places <- sample(1:12,2)

  # replace 2 random digits with the answers
  vec[places[1]] <- answ
  vec[places[2]] <- tansw

  # check the combination and stop if only 1 combination is possible
  # 1 combination is a count of 2 since selected numbers are non unique
  stopit <- matcheck(vec)
  if (stopit == 2) {
```

```
    break
  }
}

# turn the vector into a matrix
task <- matrix(vec, 4, 3)

# mark the answers in a solution (can be skipped)
vec[places[1]] <- paste0(">", vec[places[1]], "<")
vec[places[2]] <- paste0(">", vec[places[2]], "<")
solution <- matrix(vec, 4, 3)

# print the task and the solution
write.table(task, sep = ",",
            row.names=F, col.names=F, quote = F)
write.table(solution, sep = ",",
            row.names=F, col.names=F, quote = F)


######################### Unsolvable matrices for cheating
#########################

repeat{
  ##Loop to make a new vector of numbers until NO combination of numbers
adds to 10

  # create a vector of 12 random digits rounded on 2 digits
  vec <- round(runif(12, 0, 10),2)

  # check the answers and stop if no combination is possible
  stopit <- matcheck(vec)
  if (stopit == 0) {
    print(stopit)
    break
  }
}

# turn the string into a matrix
cheat <- matrix(vec, 4, 3)

# print cheat matrix
write.table(cheat, sep = ",",
            row.names=F, col.names=F, quote = F)
```

Appendix C

```
#Data analysis for thesis "Dark Tetrad of Personality and Lying in a
Cheating Task"

#Used packages
library(haven) #read_spss
library(ggplot2) #plots
library(jtools) #apa theme for plots
library(ggpubr) #qqplots
library(car) #levenes test
library(tidyr) #logistic regression
library(dplyr) #logistic regression
library(rms) #logistic regression
library(AER) #dispersion test


#Data
empdata <- read_spss("EMPdata_schoon.sav")
empdata$Gender <- factor(empdata$Gender, levels = c(1,2,3,4), labels =
c("Male", "Female", "Non-binary/third gender", "Prefer not to say"))
#recode into factor
empdata$QMatrix <- as.numeric(empdata$QMatrix)
empdata$mach <- rowMeans(empdata[9:15])
empdata$nar <- rowMeans(empdata[16:22])
empdata$psy <- rowMeans(empdata[23:29])
empdata$sad <- rowMeans(empdata[30:36])


#Descriptives
table(empdata$UserLanguage)
mean(empdata$Age, na.rm = T)
sd(empdata$Age, na.rm = T)

table(empdata$Gender)
table(empdata$Gender)/nrow(empdata)
table(empdata$source)

t.age <- table(empdata$Age) #frequency of age
sum(t.age[14:54])/sum(t.age) # %age > 30


#RQ1 SD4 Scores and Age
colSums(is.na(empdata)) #no missing data on sd4 questions but missing data
in gender and age
agedata <- empdata[!is.na(empdata$Age),]

plot(agedata$Age, agedata$mach)
plot(agedata$Age, agedata$nar)
plot(agedata$Age, agedata$psy)
plot(agedata$Age, agedata$sad)

ggqqplot(agedata$Age) #non normally distributed
```

```
ggqqplot(agedata$mach)
ggqqplot(agedata$nar)
ggqqplot(agedata$psy)
ggqqplot(agedata$mach)

cor.test(agedata$Age, agedata$mach, method = "spearman")
cor.test(agedata$Age, agedata$nar, method = "spearman")
cor.test(agedata$Age, agedata$psy, method = "spearman")
cor.test(agedata$Age, agedata$sad, method = "spearman")


#RQ2 SD4 Scores and Fender
genderdata <- subset(empdata, Gender == "Male" | Gender == "Female") #only
male and female

#t-tests

leveneTest(genderdata$mach, genderdata$Gender)
t.test(genderdata$mach ~ genderdata$Gender, var.equal = T)

leveneTest(genderdata$nar, genderdata$Gender)
t.test(genderdata$nar ~ genderdata$Gender, var.equal = T)

leveneTest(genderdata$psy, genderdata$Gender)
t.test(genderdata$psy ~ genderdata$Gender, var.equal = F)

leveneTest(genderdata$sad, genderdata$Gender)
t.test(genderdata$sad ~ genderdata$Gender, var.equal = F)

#subscale mean standard deviation by gender
tapply(genderdata$mach, genderdata$Gender, sd)
tapply(genderdata$nar, genderdata$Gender, sd)
tapply(genderdata$psy, genderdata$Gender, sd)
tapply(genderdata$sad, genderdata$Gender, sd)


##Questions with the matrix task

#### RQ3 lying freq ####
p <- ggplot(empdata, aes(x = QMatrix))
p + geom_histogram(bins = 20, color="black", fill="white") + theme_apa() +
  stat_bin(binwidth=1, geom="text", aes(label=..count..), vjust=-1.5) +
  scale_x_continuous(breaks = c(1:20))

table(empdata$QMatrix)
sum(empdata$QMatrix > 13)/nrow(empdata)
sum(table(empdata$QMatrix)[1:13])
sum(table(empdata$QMatrix)[14:19])

#### RQ4 Predicting truth or lie with sd4 ####
matdata <- subset(empdata, !is.na(QMatrix))
matdata$lies <- factor(ifelse(matdata$QMatrix > 13, 1, 0), labels =
c("truth", "lie")) #convert into factor
```

```
#Logistic regression
log.modelvars <- matdata[ c("Age", "Gender", "mach", "nar", "psy", "sad",
"lies") ] #select data for model

if(any(colnames(log.modelvars) == "Gender")){
  log.modelvars <- subset(log.modelvars, Gender == "Male" | Gender ==
"Female")
  log.modelvars <- droplevels(log.modelvars)
}

# "Age", "Gender", "mach", "nar", "psy", "sad", "lies"

colSums(is.na(log.modelvars))
log.modelvars <- na.omit(log.modelvars)
log.modelemp <- glm(lies ~ mach+nar+psy+sad, data = log.modelvars, family =
binomial)
summary(log.modelemp)
lrm(lies ~ mach+nar+psy+sad, data = log.modelvars)
summary(log.modelemp)$coef[,1] / summary(log.modelemp)$coef[,2]

#2nd model
#sepmodelvars <- matdata[ c("Age", "Gender", "mach", "nar", "psy", "sad",
"lies") ] #select data for model
sepmodel <- glm(lies ~ ., data = log.modelvars, family = binomial)
summary(sepmodel)

#compare models
anova(log.modelemp, sepmodel, test = "Chisq")
with(summary(log.modelemp), 1 - deviance/null.deviance)

# Partial regression plots
mydata <- select_if(log.modelvars, is.numeric)
predictors <- colnames(mydata)
probs <- predict(log.modelemp, type = "response")

mydata <- mydata %>%
  mutate(logit = log(probs/(1-probs))) %>%
  gather(key = "predictors", value = "predictor.value", -logit)

ggplot(mydata, aes(logit, predictor.value))+
  geom_point(size = 0.4) +
  geom_smooth(method = "loess") +
  theme_apa() +
  facet_wrap(~predictors, scales = "free_y")

vif(log.modelemp)

#RQ4 Predicting lying freq with sd4

#Poisson regression
matdata$liefreq <- matdata$QMatrix-13
matdata$liefreq[matdata$liefreq < 0] <- 0
reg.modelvars <- subset(matdata, liefreq > 0)
```

```
reg.modelvars <- subset(reg.modelvars, Gender == "Male" | Gender ==
"Female") #only male and female
reg.modelvars <- reg.modelvars[ c( "Age", "Gender", "mach", "nar", "psy",
"sad", "liefreq") ] #select data for model
colSums(is.na(reg.modelvars))
reg.modelvars <- na.omit(reg.modelvars)
#reg.modelvars$liefreq <- (reg.modelvars$liefreq)^2

reg.modelemp <- glm(liefreq ~ ., data = reg.modelvars, family = "poisson")
summary(reg.modelemp)
reg.modelemp2 <- glm(liefreq ~ mach+nar+psy+sad, data = reg.modelvars,
family = "poisson")
summary(reg.modelemp2)

dispersiontest(reg.modelemp)
pairs(reg.modelvars[,2:6])

anova(reg.modelemp, reg.modelemp2, test = "Chisq")

leveneTest(reg.modelvars$liefreq, reg.modelvars$Gender)
t.test(reg.modelvars$liefreq ~ reg.modelvars$Gender, var.equal = T)
```