

Authoritarianism and Personality:
Conceptual Issues and the Role of Biased Responding

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA

BY

Steven Getman Ludeke

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Colin DeYoung, Ph.D., advisor

May 2013

Acknowledgements

I was fortunate to have several excellent mentors during my graduate career. My interactions with Colin DeYoung, my advisor, have significantly shaped my thinking on both the methods and results of psychological research. Colin is particularly unique in the passion he shows for mentorship, and his concern for both the general intellectual development and personal well-being of his students is appreciated by all who have been lucky enough to have him in this role. The support and lively engagement I received from Colin made my time as a student both stimulating and enjoyable, and I am very grateful for all of his efforts.

Wendy Johnson provided patient explanations and a critical eye that were crucial both to my early work at Minnesota and to my continuing development as a researcher. I appreciate the incredible generosity she has shown with her time, and the rigor and passion she brings to her work continue to amaze me.

I am indebted to Chris Federico for many things. Chris presents the ideas of political psychology in a way that the listener can't help but be fascinated. He helps maintain this interest among his students with the extremely detailed and considered feedback he provides on projects – feedback that clarifies both the nature of the research question and the best method to tackle that question, and in doing so makes the research interests he sparks feel tractable to address. As a result, Chris has built the liveliest research group I've ever been a part of. For anyone watching the first meeting of his lab each semester, in which students at every stage scramble to sign up for multiple opportunities to present their work, it is easy to see that he has created something special.

I am also grateful to Matt McGue for all of the assistance he has provided to me, both directly and indirectly. On the latter count, Matt has built up an incredible resource for research with the data from the Minnesota Twin Family Study. In addition, the weekly individual differences reading group he organizes has been one of the most intellectually stimulating environments I've had the good fortune to experience. More importantly, his patient and kind advice was an integral part of my development as a researcher.

Tom Bouchard and Bob Krueger are also deserving of thanks. Both of them provided access to amazing sources of data as well as guidance in analyzing and interpreting these data. I am grateful to them both.

I am also thankful for all the help provided by fellow students such as Yanna Weisberg, Michal Reifen Tagar, and Greg Perlman. In particular, Greg's generosity with his time during my first years in the program saved me countless hours and a great deal of stress. Conversations with Yanna and Michal have led to several exciting projects, and I am looking forward to more work with both of them in the future. Finally, I am grateful to Laura Sunder-Plassmann not only for the great deal of statistical assistance she provided for Study 1 of this dissertation, but also for the invaluable moral support she has provided throughout the writing of these studies.

Last but not least, I am thankful for the record-breaking weather of Minnesota's spring of 2013, whose piles of snow and torrential rains made this dissertation possible.

Abstract

In this dissertation I address two topics concerning authoritarianism: (1) the role of self-report bias in the assessment of the correlates of authoritarianism, and (2) the question of whether authoritarianism is appropriately conceptualized as a personality trait. I addressed the first topic in Studies 1 and 2. Study 1 used two samples to highlight the role of the individual's perceptions of trait desirability in predicting bias in that individual's self-reports. The individual's views of the desirability of a trait were shown to be an effective predictor of variance that remained in self-report for that trait after the variance shared with peer-reported and/or objectively-assessed levels of that trait have been removed, indicating that people were more prone to exaggerate their levels of a trait when they personally viewed that trait as desirable. In a direct comparison against socially desirable responding measures, which identify individuals who claim exaggerated levels of general patterns of traits, individual perceptions of the desirability of traits were found to predict equal or greater amounts of bias in self-report measures, depending on the trait.

As previous studies had reported that authoritarians scored highly on measures of socially desirable responding, Study 2 applied the concepts of Study 1 to identify whether and how authoritarians exaggerate trait levels in self-reports. I found that authoritarians and nonauthoritarians were prone to distinct patterns of exaggerations in self-reports, where these exaggerations were explained by the different views of trait desirability held by authoritarians versus nonauthoritarians. There is thus nothing about authoritarianism per se that was connected to a tendency to misrepresent one's true trait levels; rather, its connection with exaggeration in self-reports derived from its association with perceptions of trait desirability.

In Study 3, I addressed recent challenges to the original "trait" conception of authoritarianism in a longitudinal twin study. I found that, consistent with the results observed for other personality traits, authoritarianism was highly stable over time, and this stability was influenced primarily by genetic factors.

Table of Contents

Acknowledgment.....	i
Abstract	iii
Table of Contents	iv
List of Tables	viii
List of Figures	ix
Chapter I: General Introduction	1
Chapter II: Idiographically Desirable Responding: Individual Differences in Perceived Trait Desirability Predict Overclaiming (Study 1).....	7
II.1 Introduction	8
II.2 Sample 1	13
II.3 Method.....	14
II.3.1 Participants.....	14
II.3.2 Measures.....	14
II.3.2.1 Personality.....	14
II.3.2.2 Ideal Self.....	15
II.3.3 Analysis.....	16
II.4 Results.....	17

II.5	Discussion.....	21
II.6	Sample 2.....	22
II.7	Method.....	22
II.7.1	Participants.....	22
II.7.2	Measures.....	23
II.7.2.1	Big Five.....	23
II.7.2.2	Desirability.....	24
II.7.2.3	Intelligence.....	25
II.7.2.4	Socially desirable responding.....	25
II.7.2.5	Self-Esteem.....	26
II.7.3	Analysis.....	26
II.8	Results.....	27
II.8.1	Trait desirability's association with single-criterion constructs....	27
II.8.2	Desirability and multiple-criteria constructs.....	29
II.8.3	Socially desirable responding.....	30
II.8.4	Individual differences in IDR.....	31
II.9	Discussion.....	32
II.9.1	Desirable responding: Two approaches.....	32
II.9.2	Comparisons between the two studies.....	34

II.9.3	Considerations for future work.....	37
II.10	Conclusion.....	40
Chapter III: Not as different as we want to be: Attitudinally-consistent trait desirability leads to exaggerated associations between personality and sociopolitical attitudes (Study 2).....		
III.1	Introduction.....	43
III.2	Method.....	45
III.2.1	Participants.....	45
III.2.2	Measures.....	46
III.2.2.1	Big Five.....	46
III.2.2.2	Desirability.....	46
III.2.2.3	Intelligence.....	47
III.2.2.4	Social attitudes.....	47
III.2.3	Analysis.....	47
III.3	Results.....	49
III.4	Discussion.....	51
Chapter IV: Authoritarianism as a Personality Trait: Evidence from a Longitudinal Behavior Genetic Study (Study 3).....		
IV.1	Introduction.....	54

IV.1.1	Previous research on the stability of authoritarianism and other personality traits.....	55
IV.2	Methods.....	58
IV.2.1	Participants and measures.....	58
IV.2.1.1	Assessment 1.....	59
IV.2.1.2	Assessment 2.....	59
IV.2.2	Analysis.....	60
IV.3	Results.....	61
IV.3.1	Phenotypic stability.....	61
IV.3.2	Biometric results.....	62
IV.4	Discussion.....	63
	Tables.....	68
	Figures.....	81
	References.....	83

List of Tables

Table 1. Associations of trait desirability with personality and self-criterion residuals in student sample.....	68
Table 2. Associations of trait desirability with trait levels and self-criterion residuals in adult community sample.....	70
Table 3. Intelligence: Associations Between Desirability, Trait Levels, and Self-Criterion Residuals (SCRs).....	72
Table 4. Correlations of the BIDR scales with self-criterion residuals.....	73
Table 5. Overclaiming Predicted in Multiple Regression Using BIDR Scales and Trait Desirability.....	74
Table 6. Correlations of sociopolitical attitudes with personality, self-criterion residuals, and trait desirability ratings.....	76
Table 7. Standardized path weights and fit indices from the model in Figure 1.....	77
Table 8. Intraclass correlations in authoritarianism scores between members of twin pairs.....	78
Table 9. Genetic and environmental contributions to authoritarianism.....	79

List of Figures

Figure 1. A mediation model to identify the influence of sociopolitical attitudes on overclaiming, by controlling for true trait variance.....	81
Figure 2. A path diagram of an AE Cholesky model for Right-Wing Authoritarianism at Time 1 and Time 2.....	82

Chapter 1: General Introduction

Inspired by the then-recent atrocities of World War II, Adorno and colleagues (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950) sought to identify the characteristics of those most likely to support authoritarian regimes. The conception of Adorno and colleagues (1950) reflected the psychoanalytic tradition then dominant in psychology and was relatively expansive, containing among its nine core elements characteristics as disparate as an inclination towards mysticism and an elevated interest in detecting and preventing sexual promiscuity. Of the many revisions offered to this construct in the decades since this seminal work, those offered by Altemeyer (1981; 1988; 1996) were the most routinely employed by psychological researchers. In his conception of Right-Wing Authoritarianism (RWA), Altemeyer (1981) discarded the psychoanalytic foundation of authoritarianism and retained only three of the features identified by Adorno et al. (1950), including conventionalism, a willingness to obey established authority, and a punitive attitude towards deviants. Efforts to develop the nomological network of authoritarianism and to understand the development of authoritarian attitudes have frequently focused on the connections between authoritarianism and broad-band personality inventories such as those assessing the Big Five. In this dissertation I addressed two questions that have been raised by this research. The first question, addressed in Studies 1 and 2, concerns the role of biased responding in this literature. The second, addressed in Study 3, concerns the conceptualization of authoritarianism; though originally conceived as a personality trait, recent work has argued it is more accurately viewed as a product of personality traits interacting with

worldviews and recent experiences, and thus more variable over time and circumstance (e.g. Sibley, Wilson, & Duckitt, 2007, Van Hiel, Pandelaere, & Duriez, 2004).

Because the overwhelming majority of the research connecting authoritarianism to other personality traits (recently reviewed by Sibley & Duckitt, 2008) has relied exclusively on self-report assessments, the interpretation of results from these studies must be tempered by concern over the accuracy of the self-report measures. The need for such concern is highlighted by the large number of studies reporting positive associations between social desirability measures and a range of right-wing attitude measures, including various measures of political conservatism (Egan, 1989), traditionalism (Shaffer & Duckitt, 2013; Tellegen, 1978/1982), and authoritarianism (Altemeyer, 1998; Meston, C., Heiman, J., Trapnell, P. D., & Paulhus, D. L., 1998; Ray, 1979; Tsang & Rowatt, 2007). As will be discussed below in Study 1, social desirability measures have evolved and proliferated in the decades since their introduction, but Paulhus (2002) suggests that two primary themes can be found in these measures. The first concerns presenting oneself as a “saint” (with elevated levels of Conscientiousness, Agreeableness, and Emotional Stability); the second, as a “superhero” (with elevated levels of Extraversion, Openness, and Emotional Stability). With respect to social attitudes, it is primarily measures of the former tendency that have been associated with measures such as RWA in the above-cited studies and elsewhere. Given that authoritarians tend to report high levels of Conscientiousness in self-report surveys (Sibley & Duckitt, 2008), researchers might thus worry whether this association accurately represents personality differences between authoritarians and nonauthoritarians.

At the same time, the use and interpretation of measures of socially desirable responding has been a source of controversy, as scores on these measures include true variance in traits such as the Big Five in addition to bias (Ones, Viswesvaran, & Reiss, 1996). Accordingly, the association of authoritarianism with measures of “saint-like” self-presentation may reflect true associations of authoritarianism with “saint-like” traits, a tendency to misrepresent levels of these traits, or both. To address this question, the use of independent criteria (and not just self-report) to assess personality traits is essential. As shown below, an examination of the association of Altemeyer’s RWA scale with both self- and peer-reports of Big Five traits pointed to a more complex picture than might be inferred simply by correlating measures of socially desirable responding with RWA and other attitude constructs. Specifically, while authoritarians in our sample were found to be particularly prone to overclaim levels of Conscientiousness, there are also traits which nonauthoritarians were particularly prone to overclaim. In seeking to understand these findings, I conducted the research presented in Study 1, which highlighted an aspect of biased responding largely overlooked in previous work. Specifically, I found that participants in two samples had particularly overclaimed those traits that they personally viewed as desirable. This is in contrast to socially desirable responding measures, which (as their name suggests) identify the tendency to claim high levels of those traits viewed positively by society at large. Study 2 applied this general finding to the domain of sociopolitical attitudes: attitude measures like RWA were associated with biased responding in personality, but these associations appeared to reflect attitude-based differences in views of trait desirability rather than any process specific to attitudes. That is, the reason those low and high in authoritarianism tended to overclaim different traits

was because they tended to value different traits; two individuals who held different sociopolitical views but comparable views of trait desirability might thus be expected to show similar patterns of overclaiming.

An additional question raised by recent research in authoritarianism and personality concerns the question of the appropriate conceptualization of authoritarianism. Early formulations of authoritarianism viewed the construct as a personality trait (e.g. Adorno et al., 1950; Altemeyer, 1988), with Altemeyer arguing that “[t]he RWA scale is a personality test disguised as an attitude survey” (Altemeyer, 2009, p. 14). This attitudinal-phrasing has been the source of some significant controversy: for example, Jugert and Duckitt (p. 464, 2009) noted that “[i]t has been pointed out that the RWA items are statements of social attitudes and beliefs and that there is no evidence that they measure personality,” consistent with the conception expressed in Sibley and Duckitt (2008) that personality traits should be viewed as “purely behavioral dispositions” (p. 276). Because items assessing attitudes are present in the majority of broad-band personality measures, however, (see Pytlik Zillig, Hemenover, & Dienstbier, 2002; Werner & Pervin, 1996) it may be that this aspect of the dispute simply reflects a lack of familiarity with personality measures on the part of those challenging the trait conception of authoritarianism.

Other challenges to the trait conception of authoritarianism have focused on whether authoritarianism functions similarly to other personality traits. For example, Sibley et al. (2007) argued that authoritarianism is too malleable to be considered a personality trait. Though critiques of the trait-conception of authoritarianism have often not explicitly connected their proposals to the ideas of personality theorists or provide

clear definitional boundaries between personality traits and other kinds of psychological constructs, critics of the trait-conception often appear to view authoritarianism along the lines of what McAdams and Pals (2006) labeled “characteristic adaptations.” In comparison to traits, characteristic adaptations are more responsive to cultural and environmental influence, more subject to change over time, and have effects on behavior that are more dependent on situational context. Critics of the trait conception of authoritarianism have often suggested authoritarianism is not a trait itself but instead reflects both stable personality traits as well as the extent to which an individual currently views the world as a safe versus a dangerous place, where the hypothesized influence of this supposedly fluctuating worldview is suggested to introduce a degree of mutability to authoritarianism that is inconsistent with a conceptualization of it as a personality trait (e.g. Sibley et al., 2007). In Study 3 I reviewed the conceptual issues involved in this challenge and explored the longitudinal stability of authoritarianism over a 15-year period in a sample of twins, comparing the stability for this trait and the biometric sources of this stability to that observed for undisputed personality traits. Consistent with the “trait” conception of authoritarianism, results from this study paralleled those observed in previous work for other personality traits. Importantly, while the observed stability for authoritarianism was as high as that reported in meta-analyses of other personality traits (Roberts & DelVecchio, 2000), like them, it was also not perfect: the correlation between the two time points was .74, and corrections for attenuation due to measurement error only increased this value to .79. It thus may be that this debate derives primarily from an idiosyncratic conceptualization of personality among opponents of the trait conception of authoritarianism, as the requirement that traits be immutable (Sibley et

al., 2007; p. 358) and assessed without reference to attitudes does not appear to be met by any trait (Pytlik Zillig, Hemenover, & Dienstbier, 2002; Roberts & DelVecchio, 2000; Werner & Pervin, 1996).

Chapter II: Idiographically Desirable Responding: Individual Differences in Perceived Trait Desirability Predict Overclaiming (Study 1)

Objective

Conventional measures of self-report bias implicitly assume consistent patterns of overclaiming across individuals. We contrast this with the effects of individual differences in views of trait desirability on overclaiming, which we label Idiographically Desirable Responding (IDR).

Method

We obtained self- and peer-reports of trait levels on mixed-sex samples of undergraduates (N=352) and middle-aged community members (N=541), with an additional performance-based assessment in the latter sample.

Results

Compared to conventional measures of bias, individual differences in trait desirability ratings identified an independent and comparatively large amount of the variance in overclaiming for personality and physical attractiveness. The importance of IDR was confirmed by the replication of these results for intelligence, for which both peer-ratings and performance data were available. Individuals differed in the extent to which they rely on IDR, with these differences indexed by the correlation between views of the desirability of a given trait and the extent to which one overclaimed that trait. Individuals who were more prone to overclaim in this fashion exhibited higher self-esteem as well as higher scores on questionnaire measures of socially desirable responding.

Conclusion

Overclaiming of traits resulted both from the patterns of biases identified by conventional overclaiming measures and from individual differences in perceptions of what traits are most desirable.

II.1 Introduction

Self-report measures of personality have a number of theoretical and practical virtues (Paulhus & Vazire, 2007), and a substantial body of literature supports their predictive utility in both basic research and applied settings (e.g., Roberts et al., 2007). Nonetheless, concerns over the vulnerability of these instruments to distortion have been raised for over a half-century (Edwards, 1957). A sizeable literature has developed to identify those who might provide biased self-reports, as well as the contexts that promote bias in these measures. In their survey of the various efforts to identify biased responding in self-report, Paulhus and John (1998) argued that existing measures typically capture one or another of two distinct patterns of distortion. “Egoistic bias” refers to the tendency to claim high social and intellectual status, whereas “moralistic bias” refers to the tendency to deny socially deviant impulses and behaviors. Instruments assessing each of these tendencies measure individual differences in the degree to which a person is prone to a given manner of response. An assumption implicit to these instruments is that, at least within each category of bias (i.e., egoistic or moralistic), the *pattern* of biased response – that is, the traits that people exaggerate – will be the same from one individual to the next.

This implicit assumption may obscure an alternative source of bias. The present study investigates the role of individual differences in perceptions of trait desirability in

producing bias in self-reported trait levels. Rather than expecting misrepresentation in self-report to follow only one of a limited set of patterns (egoistic or moralistic), we additionally expect individuals to self-report higher levels of traits to the extent that they view them as desirable. The conceptions of “desirable traits” operationalized in existing measures were derived from aggregations of individual opinions on traits (e.g., Jackson & Messick, 1962); this literature appropriately refers to its topic as “socially desirable responding” (SDR) because it is the average desirability of the trait in society that is thought to provide the motivation for distortion. Consistent with this terminology, we label the phenomenon studied here “idiographically desirable responding” (IDR) to indicate that it is each individual’s personal views of the relative desirability of traits that are expected to be an importantly distinct source of misrepresentation in their self-reported personality.

While a vast, venerable, and growing literature explores SDR, investigations of anything resembling IDR have been much rarer. This may reflect a belief that few meaningful differences exist between individuals’ perceptions of the desirability of traits. This belief has some basis, as mean inter-rater correlations for ratings of trait desirability typically surpass .6 (e.g., Krueger, 1996). Nevertheless, Sinha and Krueger (1998) found indications of possible influence of desirability ratings on the overclaiming of a given trait, showing that individuals who rated a trait as more desirable than did others also tended to claim higher levels of that trait. This study provided an important first step toward examining IDR, but could not resolve the question of whether this association truly represented bias; their results are equally consistent with a tendency of individuals to rate as more desirable the traits they *actually* possess. That SDR measures have long

been shown to measure at least as much true variance as bias in self-rated traits (McCrae & Costa, 1983; Ones et al., 1996) highlights the challenges to interpreting any association between self-reported trait levels and any proposed method of identifying bias. What is needed to do so effectively is evidence from an external criterion.

Borkenau, Zaltauskas, and Leising (2009) carried out the only study we know of to use an external criterion in an exploration of IDR. This study departed from most previous work on the topic in two ways. First, instead of assessing the relative desirability of traits, it assessed perceptions of the optimal level of traits. Second, it investigated how individual differences in these perceptions predicted distortions in self-reported personality. The optimal level of traits (e.g. talkativeness, suspiciousness) was assessed on a 6-point Likert-type scale by each subject, resulting in an idiographic index labeled the Perceived Optimal Trait Level (POTL). Significant associations observed between these ratings and peer-reported trait levels indicated that differences in views of optimal trait levels were, in fact, associated with differences in actual personality. However, ratings of optimal trait levels also predicted variance in self-report trait levels not shared with peer-report, indicating that POTL also represents a source of bias in self-report.

Borkenau et al.'s (2009) use of peer-reported personality levels as a criterion to indicate "true" trait levels is consistent with recommended practice in the field (Paulhus & John, 1998) and has considerable empirical and conceptual support (Connelly & Ones, 2010). At the same time, self- and peer-reports of personality are consistently shown to provide mutual incremental improvement in predictive validity (Connelly & Ones, 2010) indicating that at least some of the variance not shared between self- and peer-report is reflective of true trait levels. Having established POTL as an indicator of true trait

variance, the results of Borkenau et al. (2009) leave open the possibility that some (or even all) of the variance found in self- but not peer-report that is shared with POTL was in fact true trait variance. For this reason, the use of additional measures of “true” trait levels would provide a more conclusive demonstration of the existence and import of IDR. While objective measures are infrequently used for traits such as extraversion and agreeableness, both SDR and IDR have obvious potential connections to attributes for which objective assessments are more widely employed, such as intelligence. Indeed, existing measures of SDR have already been shown to predict overclaiming of intelligence when compared to performance on an IQ test (Paulhus & John, 1998). The inclusion of such measures in addition to peer report would thus provide a more conclusive demonstration of IDR. For this reason, we included a performance test of intelligence in Sample 2.

To quantify the extent of overclaiming present in self-reports, Paulhus and John (1998) suggested the use of self-criterion residuals (SCR), in which bias is captured by the residual variance remaining after regressing a self-report measure on a criterion measure of the same variable. After the variance shared with the criterion has thereby been removed from the self-report, any variance due to biased responding is thought to remain in the residual. SCRs have been repeatedly used in studies of SDR measures (Paulhus et al., 2003; Pauls & Crost, 2003), and the relative effectiveness of SDR and IDR at identifying these residuals is therefore of interest. Further, such a comparison is needed to investigate the independence of SDR and IDR. The importance of IDR would be supported if it could predict variance in SCRs over and above measures of SDR.

Both SDR and IDR can be operationalized in several ways. A range of instruments assessing SDR have been developed in the past 50 years (Paulhus, 2002). With scales to assess both egoistic and moralistic bias, the Balanced Inventory of Desirable Responding (Paulhus, 1991) is among the most used in recent years. Procedures to assess IDR are less well established. The optimal-trait-level approach employed by Borkenau et al. (2009) has much to recommend it, including the important insight that the desirability of most personality traits is not a strictly linear function. However, this approach does not have the ability to compare differences in the *degree* of desirability of different traits: two traits may be matched perfectly in perceived optimal trait level but differ markedly in the desirability of attaining that level. As SDR measures were typically developed with an eye toward the relative desirability of a given trait (e.g., Edwards, 1953), an IDR index based on individual differences in the perceived desirability of a trait may be the most direct method of comparison between SDR and IDR.

We have thus far stressed how individual differences in perceived trait desirability may lead to individual differences in which traits a given individual overclaims. We expect these effects to be evident among the population as a whole, with those who view a trait as particularly desirable the most likely to have overclaimed that trait. But there may also be individual differences in this tendency: that is, some individuals may be particularly prone to overclaim the traits they personally view as desirable.

Differences of this nature are plausibly connected with several extant constructs. For example, those who are particularly prone to overclaim the traits they personally view as desirable (IDR) may also tend to overclaim traits viewed as desirable by society

at large (SDR). This would suggest a general willingness to overclaim desirable traits, and would be supported if high scorers on SDR measures were also more prone to IDR.¹ Further, just as those high in SDR tend to exhibit high self-esteem (Mar, DeYoung, Higgins, & Peterson, 2006), individual differences in IDR may show the same relationship, as those with a positive view of themselves may be particularly prone to believe they have the traits they view as desirable.

II.2 Sample 1

We present results from two samples that show how individual differences in trait desirability connected both to true differences in personality as well as to bias in self-reports. Sample 1 consisted of members of romantic couples who provided personality ratings both for themselves and their partner in addition to completing a measure of the desirability of those traits. We examined correlations between trait desirability ratings and partner ratings to test whether trait desirability is associated with accuracy in self-report. To test whether trait desirability is associated with bias in self-report, we examined the correlations between trait desirability and SCRs obtained by regressing

¹ The interactions between individual and societal perceptions of trait desirability are likely quite complex, with different effects in different individuals and in different contexts. For an illustration of the idea at its simplest, consider four traits: Trait A is viewed as neutral by both the individual and society, trait B is viewed as desirable by the individual but neutral by society, trait C as neutral by the individual but desirable by society, and trait D as desirable by both the individual and society. Those low in both IDR and SDR will report all traits accurately; those high in IDR but not SDR will overclaim A and B but report C and D accurately; those high in SDR but not IDR will overclaim C and D but report A and B accurately; those high in both SDR and IDR will report A accurately, overclaim D the most, and will overclaim B and C equally.

self-ratings on partner ratings. Individual differences in the tendency to be influenced by IDR were also examined.

II.3 Method

II.3.1 Participants

We recruited 187 heterosexual couples via advertisements on the psychology department research website and flyers posted throughout the campus and surrounding neighborhood of a large public university. All recruitment materials specified that couples had to be at least 18 years of age and dating for a minimum of three months and a maximum of one year. Of the 176 couples to provide complete data, the average relationship length was 10.5 months ($SD= 13.04$, median = 7). The mean age of the women was 20.5 years ($SD=4.28$), and the mean age of the men was 21.46 years ($SD=4.77$). The majority of participants were Caucasian (83.8%), 11.4% were Asian, 2.2% were Hispanic, less than 1% Black, and 4% other or declined to provide ethnicity. Ninety-five percent of participants reported dating their partner exclusively; 3 couples were engaged. Each participant was compensated with \$25 or ten extra credit points to use toward a psychology course. Both partners of each couple independently completed online measures.

II.3.2 Measures

II.3.2.1 Personality

Personality (both self- and peer-ratings) was assessed using a shortened version of the Big Five Aspect Scales (BFAS; DeYoung et al., 2007). The BFAS allows for assessment of personality traits both at the level of the Big Five and at the sublevel of the two aspects

within each domain. The shortened version of the scale used in this research included four items assessing each aspect of the Big Five instead of the full ten items, for a total of 40 items. The four items were selected based on validity evidenced in multiple samples of undergraduate students previously collected. Each item was correlated with overall scores on the Big Five and aspects as measured by the full 100-item BFAS. Then, correlations between each possible four-item scale with overall scores on the full BFAS were calculated. The four-item scales that showed the highest correlations ($r_s = .82$ to $.94$) with the full BFAS scales and evidenced the greatest reliability ($\alpha_s = .64$ to $.81$) were chosen. The short BFAS was repeated with different instructions in order to assess both ratings of one's own personality (self ratings) and ratings of one's partner's personality (peer ratings). Participants rated each item in terms of how well the statement described the self or the partner using a 5-point scale on which 1 was labeled *strongly disagree* and 5 was labeled *strongly agree*.

II.3.2.2 Ideal Self

The short BFAS was repeated with different instructions in order to assess one's ideal personality. Instructions read:

Here are a number of characteristics that may or may not describe your **current ideal self**. For example, do you agree that you *ideally* would seldom feel blue? Please fill in the number that best indicates the extent to which you agree or disagree that each statement reflects your ideal self. Be as honest as possible, but rely on your initial feeling and do not think too much about each item.

Participants rated each item in terms of how well the statement described the ideal self on a 5-point scale on which 1 was labeled *strongly disagree* and 5 was labeled *strongly agree*. A selection of 5 on “seldom feel blue” indicates that they strongly agree that their ideal self would seldom feel blue, and thus represents the desirability of this trait. Desirability ratings for each of the Big Five trait were then computed by taking the mean of each of the relevant ratings.

II.3.3 Analysis

Two different self-criterion residuals (SCRs) were obtained using peer report as the criterion measure. The first assessed differences between participants in the extent to which they over- or under-claimed a single trait (e.g. Extraversion), relative to other participants in the sample. We label this SCR a “within-trait self-criterion residual” (WT-SCR). These SCRs are an individual difference measure appropriate for comparisons across participants—for example, testing the association between the desirability of a trait and the extent of overstatement of that trait in the sample as a whole. WT-SCRs were obtained by regressing the self-report scores for a given trait on the peer-report scores *across all participants*, with the standardized residuals saved as a new variable and used for later analysis. This procedure removes the variance in self-reports that is shared with peer-reported trait levels, leaving in the residual any tendency to represent trait levels as higher or lower than what would be predicted based on partners’ ratings. Thus, an individual with a large, positive WT-SCR for Extraversion overclaimed that trait to a greater extent than did most other participants.

The second SCR assessed the extent to which a given participant over- or under-claimed one trait relative to other traits. We label this SCR a “within-participant self-

criterion residual” (WP-SCR). These SCRs are necessary for analyzing trends within an individual, such as the extent to which his or her misrepresentations of trait levels are predicted by personal views of trait desirability. WP-SCRs were obtained by regressing self-report scores for all traits on peer-report scores for all traits in a regression run *across all traits within a single participant*. Standardized residuals were saved as a new variable and used for later analysis. As with WT-SCRs, this procedure removes the variance in self-report of a trait that is shared with peer-reported trait levels, leaving in the residual any tendency of the individual to represent the level of a given trait differently than would be predicted by the partner’s ratings of that trait. WT-SCRs and WP-SCRs differ in that the former uses variance in self-reported levels of a given trait by the entire sample, while the latter uses variance in self-reported levels of all traits within a given individual. The WP-SCRs can thus be described as an index of the extent to which an individual overclaims each trait relative to the other traits, whereas the WT-SCRs can be described as an index of the extent to which an individual overclaims each trait relative to other individuals.

Though our sample size provides a good deal of power for analyzing across-participant comparisons for each trait (WT-SCRs), the power for within-subject comparisons (WP-SCRs) is determined by the number of traits being analyzed. The number of traits assessed is much smaller than the number of participants, and thus it is more difficult to achieve adequate power in analysis of WP-SCRs. Therefore, although WT-SCRs were computed using aggregated Big Five traits, WP-SCRs were computed using each of the 40 items assessed in the short-form BFAS. Thus, a participant with a

large, positive WP-SCR for “Have a lot of fun” overclaimed that item more than he or she overclaimed other items.

Finally, although our construct of interest concerns the tendency to exaggerate traits because of their idiographic desirability, differences in the way people use Likert scales may also influence the associations of trait desirability with self-reports and thus with SCRs. For example, individual differences in tendencies such as extreme responding (in which individuals predominantly use the ends of the Likert scale) will produce an increase in associations between all self-report measures. This would exaggerate the role of idiographic desirability in inducing overclaiming, and a more conservative test of IDR requires eliminating effects due to extreme responding. To remove any such effects from desirability ratings we created residuals for each of the Big Five, predicting each trait in turn with the remaining four. These residualized desirability ratings thus contained the desirability a given individual assigned to that Big Five trait after removing any variance shared with the desirability ratings of other Big Five traits, including any variance due to extreme responding. It was these residualized desirability ratings that were then used in the analyses below; supplementary analyses using nonresidualized desirability ratings provided substantively the same pattern of associations that is reported below, though with slightly larger values.

II.4 Results

Desirability ratings for each trait were strongly correlated with self-report ratings for the same trait. As shown in Table 1, desirability ratings were highly correlated with their corresponding self-rating across the Big Five (e.g., the association of

Openness/Intellect desirability ratings and Openness/Intellect self-ratings was .64), with a mean association of .60. Further, these associations were highly specific: the absolute value of every cross-trait correlation (e.g. between Openness/Intellect desirability and Emotional Stability self-ratings, $r = -.06$) was below .20, and the mean of the absolute values of all cross-trait correlations was .07.

Insert Table 1 about here

The association of desirability ratings with self-ratings of trait levels may indicate the presence of a true association in which individuals value the traits they actually possess, but it is also consistent with an idiographic response tendency in which people over-report the presence of traits they value. A third option is that both accuracy and bias in self-reported personality are indicated by these associations. To address the question of accuracy, we explored the association between the target's ratings of trait desirability and peer ratings of the target's personality. As with self-reported personality, Table 1 shows that highly significant associations were generally found between desirability levels and peer-reported trait levels, with a mean association across the Big Five of .31. These associations were also highly specific, with no cross-trait association above .20 in absolute value and a mean absolute cross-trait association of .04.

There are thus clear indications that trait desirability ratings are associated with actual personality levels. Nevertheless, the magnitudes of the associations with peer-reported personality were considerably smaller (mean of .31) than they had been with self-reported personality (mean of .60). It is thus worth exploring possible associations

with bias – that is, whether those rating a trait as particularly desirable tend to systematically over-report their levels of the trait in question. The associations between trait desirability and WT-SCRs shown in Table 1 indicate that an effect of this nature is present, with strong associations for all traits and a mean association of .52. These effects are also highly specific, with only one cross-trait correlation above .20 in absolute value and a mean absolute cross-trait association of .07.

The substantial associations between trait desirability and WT-SCRs indicate that someone who viewed a given trait as highly desirable was more likely to overclaim that trait than was someone who viewed the trait as less desirable. (For negatively valenced traits, the converse holds: someone who viewed a trait as particularly undesirable was more likely to under-claim that trait than was someone who viewed the trait as less undesirable.)

To explore whether participants differed in the degree to which their views of trait desirability influenced their tendency to over- or under-claim a particular trait requires analyzing the associations between desirability and SCR computed within each participant (WP-SCRs). The mean and median correlations between desirability ratings and WP-SCRs were .46 and .49, indicating that the average participant's tendency to overclaim traits was strongly predicted by his or her perception of the desirability of those traits. The tendency to over- or under-claim traits based on desirability appears to be reliably assessed (the Spearman-Brown adjusted split-half reliability of odd- and even-

numbered items was .77) and to vary considerably between individuals (range: -.40 to .88, $SD = .24$).²

II.5 Discussion

The correlations between people's self-reported ideal characteristics and the ratings provided by their partners of their actual characteristics indicate that, to a significant extent, people truly possessed the characteristics that they viewed as valuable. At the same time, the associations between self-reported ideal characteristics and self-rated actual personality were considerably stronger, indicating that although people may tend to have the characteristics they value, this association may be stronger in their own perceptions than in reality. WT-SCRs index the remaining variance in a self-report after the variance shared with the criterion measure is removed, and would thus capture any bias that was present in self-report but absent from the criterion. The strong correlation between self-rated ideal personality and these residuals indicated that bias of this nature was present. We believe this pattern is most consistent with a tendency to value the traits one has (shown by the association between peer-ratings and self-rated ideals) as well as a tendency to imagine oneself as having slightly more desirable traits than one truly does. Using WP-SCRs, we further demonstrated that while the typical participant tended to overclaim traits as a function of how desirable he perceived that trait to be, considerable variance existed between individuals in this tendency.

One limitation of Sample 1 is that only the partner's report of the target's personality was available as an indicator of true trait levels. Previous work has indicated

² The gaps between self- and peer-reports can also be obtained using difference scores, and a highly comparable pattern of results was obtained when these were used in place of both WT-SCRs and WP-SCRs.

that partners share similar values and may be prone to see each other in an overly positive light (Murray, Holmes, Bellavia, Griffin, & Dolderman, 2002). To the extent that this is the case, self-reported values of traits would show inflated associations with partner-reported trait levels and suppressed associations with self-criterion residuals obtained using partner report as the criterion. The use of an objective criterion measure is important for confirming the reality of the connection between trait valuation and true trait levels.

II.6 Sample 2

In Sample 2 we looked to replicate and extend the results found with Sample 1. Whereas the latter had only one peer rater of trait levels, each trait in Sample 2 was assessed with multiple raters, providing a more accurate picture of trait levels (Connelly & Ones, 2010). For a particularly rigorous test of IDR's ability to identify bias in self-report we used an additional measure for one trait, assessing intelligence not only by peer report but also with an objective test. We also tested whether IDR's effects can be seen for traits outside of the mental domain by looking at physical attractiveness.

To establish the relative importance of bias due to IDR, its ability to predict overclaiming was compared against an established SDR measure. Finally, we explored the nature of individual differences in the tendency to engage in IDR. First we hypothesized a connection between IDR and SDR, which would indicate the existence of a general willingness to claim desirable traits. Second, we hypothesized that those with a positive view of themselves would be more likely to believe they had the traits they viewed as most desirable, even to the point of overclaiming those traits.

II.7 Method

II.7.1 Participants

Participants in this study were members of the Eugene-Springfield Community Sample (ESCS), a group that was recruited by mail from lists of homeowners in the Eugene-Springfield area. ESCS participants agreed to complete questionnaires delivered by mail over a period of 10 years beginning in 1994 in exchange for cash. In line with the community from which the sample was drawn, most participants identified as White (97%), with 1% or less identified as Hispanic, Asian American, Native American, or declining to report their ethnicity. The sample is primarily middle aged (Mean age = 51 years, $SD = 12.36$, range = 18-80), and reported a range of educational attainment, with a median of two years of post-secondary education completed. The measures used in the present study were assessed over a period of several years, reducing the total N available for analysis. We used all participants for whom data was available for self-report on personality, trait desirability, and two or more peer-reports of personality. After excluding two participants whose responses indicated obvious inattention (all 97 trait desirability ratings were marked as “neither desirable nor undesirable), our sample for the primary analyses below was 541 (58% female), though a somewhat smaller portion of this pool had completed the additional measures of intelligence, SDR, and self-esteem described below. The 1,520 individuals providing peer-report data were 61% male, ranging in age from 6 to 94 (Mean age = 48 years, $sd = 17.81$). Previous work on informant reports in this sample (DeYoung, 2006) showed no effect from removing younger raters from the sample, so the present analyses were conducted using all available peer reports.

II.7.2 Measures

II.7.2.1 Big Five

Big Five personality traits were assessed with a mailing of the 44-item Big Five Inventory (BFI; John & Srivastava, 1999) and Saucier's (1994a) 40-item Mini-Markers (SMM), with two items added to each instrument assessing physical attractiveness. These measures were also completed independently by two or three peers who knew the participants well and were asked to rate them. Both instruments were completed in Fall of 1998 using a Likert-type scale with answers ranging from 1 (strongly disagree) to 5 (strongly agree). Scores for Big Five personality and physical attractiveness for both self- and peer-report were obtained by taking the mean of all items for each trait from both the BFI and SMM, yielding alphas between .84 and .94. As significant correlations are often found between IQ and Openness/Intellect (DeYoung, 2011), it is no surprise that many of the Openness/Intellect items have high face validity for assessing intelligence (e.g. "unintelligent," "perceptive"). A subset of Openness/Intellect items judged by the authors to be most directly relevant to intelligence were selected, including Complex, Deep, Intellectual, Philosophical, Unintelligent, "Is ingenious, a deep thinker," and "Likes to reflect, play with ideas." Post-hoc analyses confirmed that mean peer-report for each item correlated at .25 or above with objective intelligence, and the resulting seven-item scale had alphas for self- and peer-report of .84 and .87, respectively.

II.7.2.2 Desirability

To assess individual differences in beliefs about the value of certain traits, participants were asked to rate "how desirable or undesirable you feel it is for others to be or act this way" for a list of 97 characteristics constituted primarily by personality traits. Responses were provided in a mailing in summer of 2001, with participants rating each

characteristic on a 9-point Likert-type scale anchored by “very undesirable” and “very desirable.” Twenty-five of these adjectives are included in Saucier’s Mini-Markers, but an additional 54 were shown to be associated with a Big Five trait in Saucier and Goldberg’s (1996) analysis of the factor structure of personality adjectives. Scores for the desirability of Big Five traits were obtained by taking the mean of the adjectives associated in those reports with a given trait, yielding scales with the following properties: Openness/Intellect (11 items, $\alpha = .78$), Conscientiousness (19 items, $\alpha = .85$), Extraversion (17 items, $\alpha = .68$), Agreeableness (21 items, $\alpha = .87$), and Emotional Stability (11 items, $\alpha = .61$). The desirability of physical attractiveness is indicated by a composite of the desirability of “attractive” and “sexy,” which correlate .46. Finally, several of the items have high face validity for assessing intelligence: the authors identified Complex, Insightful, Perceptive, Philosophical, and Unintelligent, with the resulting scale yielding an alpha of .66. Each of these came from the Openness/Intellect scale. Sixteen remaining adjectives that were assessed in this sample did not have loadings above .30 on any Big Five trait in Saucier and Goldberg’s (1996) analysis, and were thus not incorporated in the present study.

II.7.2.3 Intelligence

Participants in the ESCS completed Cattell’s 16 Personality Factor Questionnaire (16PF; Conn & Rieke, 1994) in 1996. Factor B of this instrument is a 15-item performance test of intelligence consisting of knowledge and reasoning problems with multiple choice answers. Previous work has found the scale to correlate .57 with Full Scale IQ from the Wechsler Adult Intelligence Scale – Revised (Wechsler, 1981), supporting its use as an intelligence assessment (Abel & Brown, 1998).

II.7.2.4 Socially desirable responding

Tendencies towards socially desirable responding (SDR) were assessed in the ESCS in 1998 with the Balanced Inventory of Desirable Responding (BIDR; Paulhus 1991). In this edition of the BIDR, two different dimensions of SDR are assessed: Impression Management (IM) and Self-Deceptive Enhancement (SDE). IM captures moralistic denial of socially-deviant behaviors, while SDE measures egoistic overconfidence and claims to superiority. The alphas for these scales (IM = .82; SDE = .68) were comparable to that found in previous work (Li & Bagger, 2007).

II.7.2.5 Self-Esteem

The Rosenberg Self-Esteem scale (Rosenberg, 1965) measures global feelings of self-worth, and was administered in 1998 to ESCS participants with a 5-point Likert-type scale. The alpha reliability for this measure was .84 in the present sample.

II.7.3 Analysis

Based on the rationale and methods described for Sample 1, we obtained both WT-SCRs and WP-SCRs using peer report data for Big Five and attractiveness. For intelligence, we used objective performance as assessed by Cattell's Factor B in addition to peer report, as described further below. As noted above, WT-SCRs assess individual differences in the tendency to self-report trait levels differently than would be indicated by the criterion. In contrast, WP-SCRs identify differences within a person in the traits they most over- or under-claim.

The desirability ratings used in this study were also adjusted in the same fashion as with Sample 1 in order to reduce the role of response biases such as extreme-responding (Paulhus & Vazire, 2007). We predicted the desirability for each trait from

the desirability of the others in a regression and saved the residual, which thus represents the desirability of a given trait after removing the variance shared with the desirability ratings of other traits. (Because our assessment of intelligence's desirability derives entirely from Openness/Intellect items it was not entered into the regression for other traits; its residual was thus obtained using all traits but Openness/Intellect as predictors.) It was these residualized desirability ratings that were then used in the analyses below.

II.8 Results

II.8.1 Trait desirability's association with single-criterion constructs

Replicating the results found with Sample 1, desirability ratings for the Big Five and attractiveness were correlated with self-report ratings for the same. As shown in Table 2, the associations between desirability ratings and their corresponding self-ratings were all highly significant, with a mean across the Big Five of .35. That is, the more an individual believed a given trait was desirable, the more of that trait he or she claimed to have. As with Sample 1, these associations were highly specific, with a mean absolute value of Big Five cross-trait correlations of .05. Two such associations were significant at $p < .05$ (uncorrected for multiple testing) in both samples: those who viewed Openness/Intellect as desirable tended to report lower levels of Conscientiousness, while those who viewed Conscientiousness as desirable tended to report lower levels of Openness/Intellect. Outside the Big Five, beliefs about the desirability of attractiveness were modestly correlated with self-reported levels.

Insert Table 2 about here

The association between self-ratings of desirability and peer-ratings of the target's personality are shown in Table 2 and were also highly significant, with a mean association across the Big Five of .24. The relatively low association found for Emotional Stability is consistent with previous work showing relatively low peer accuracy for this trait (Vazire, 2010), as any true variance in personality associated with desirability ratings for Emotional Stability may be harder to detect because peer-reports are a less effective criterion for this trait than they are for others. It may also be influenced by the relatively low alpha (.61) noted above for our scale assessing the desirability of Emotional Stability. The mean absolute value of all Big Five cross-trait associations was .05, indicating these associations were highly specific. Outside of the Big Five, peer reports of the target's attractiveness were uncorrelated with the target's perception of the desirability of attractiveness (.06, $p > .1$), though targets who assigned high desirability to Agreeableness were viewed as slightly more attractive ($r = .18, p < .001$). Interpretation of these results is difficult because of the low reliability of the two-item scale used to indicate the desirability of this trait.

As with Sample 1, trait desirability ratings did not simply indicate true trait levels (indexed by the various criteria), but were also associated with systematic overclaiming of the trait levels involved. Table 2 shows the associations between trait desirability and WT-SCRs, with significant associations for all traits and a mean across the Big Five of .26. These associations were highly specific, with a mean absolute value of the cross-trait correlations of .03. One such association was significant at $p < .05$ (uncorrected for multiple testing) in both samples, with those viewing Emotional Stability as desirable

tending to understate their level of Agreeableness. Outside of the Big Five, those who viewed physical attractiveness as desirable were more likely to overclaim the trait.

This sample was more heterogeneous than Sample 1, but controlling for participants' age, sex, and education produced no substantive changes to the pattern of results presented in Table 2.

II.8.2 Desirability and multiple-criteria constructs

As discussed above, SCRs represent any variance in self-reported personality not shared with the criterion, and as such have been proposed as an indicator of the extent to which an individual over- or under-claims a trait (Paulhus & John, 1998). Nevertheless, as self-reports have been consistently shown to incrementally improve the predictive validity of peer report (Connelly & Ones, 2010), some of the variance in self report that accurately reflects trait levels may not be shared with peer reports. To the extent that this is the case, some of the variance in SCRs represents true variance in trait levels. This represents a potential challenge to the interpretation of the association between desirability and SCRs as a demonstration for IDR. One way to surmount this challenge is to compare peer-reports to a more objective criterion, which we did in the present context using the trait of intelligence, for which we have a performance based measure as well as peer-report levels.

 Insert Table 3 about here

As shown in Table 3, intelligence behaved similarly to the traits discussed above: perceptions of the desirability of intelligence were associated with self-reported levels,

true trait levels, and SCRs. Consistent with expectations, peer-reports did not remove all true trait variance from self-reports: a WT-SCR created using peer-report was modestly but significantly associated with performance on our intelligence measure ($r = .16, p < .001$). At the same time, trait desirability is significantly associated with a dual-criterion WT-SCR (which has had variance self-report shares with peer-report and with objective intelligence); in fact, this correlation is not smaller ($p > .1$) than was the correlation between desirability and the peer-report WT-SCR. Thus, the effects of IDR were not only present even when using multiple indicators of trait levels, but the addition of an objective measure of trait levels did not provide a significant reduction in the ability of trait desirability to predict overclaiming. The heavy reliance on peer-reported trait levels as criteria in studies of overclaiming (Paulhus et al., 2003; Pauls & Crost, 2003) thus appears to be well-justified, particularly given the difficulty of performance-based measurement of Big Five trait levels (Furr & Funder, 2007).

II.8.3 Socially Desirable Responding

The results reported above indicate that individual differences in perceptions of trait desirability were associated both with true differences in personality and with a tendency to overclaim traits viewed as desirable. Consistent with previous studies (Pauls & Stemmler, 2003) we found that two existing scales measuring socially desirable responding (Impression Management and Self-Deceptive Enhancement; Paulhus, 1991) were associated with bias in personality self-report. As shown in Table 4, Impression Management was associated with a tendency to overclaim these same characteristics as measured by WT-SCRs. Self-Deceptive Enhancement predicted a tendency to overclaim these traits as well as both Openness/Intellect and Intelligence. (We used the intelligence

residual which removed the variance from both peer-report and objective intelligence for the analyses presented in Tables 4 and 5.)

 Insert Tables 4 and 5 about here

Results from multiple regressions in which overclaiming of a given trait was predicted simultaneously by the desirability of that trait as well as IM and SDE are presented in Table 5. Desirability ratings remained significant predictors for their respective traits, and were superior to IM and SDE in predicting overclaiming for Openness/Intellect, Extraversion, Agreeableness and intelligence. To confirm the independence of IDR from SDR, we ran two stepwise multiple regressions for each trait. The first of these entered the BIDR measures first and then tested for improvements in R^2 after entering trait desirability ratings. The second entered desirability ratings first and then tested for improvements in R^2 after entering BIDR scores. In both of these analyses, all associations identified as significant in Table 5 produced significant improvements in R^2 , confirming the independence of SDR and IDR.

II.8.4 Individual differences in IDR

Participants in Sample 1 varied in the extent to which desirability influenced a given participant's level of over- or under-claiming trait levels. To replicate this result we analyzed the associations between WP-SCRs with trait desirability levels within participants in Sample 2. Within-subjects comparisons would not be adequately powered at the level of the Big Five, but as noted above, 25 of the 44 adjectives assessed in Saucier's Mini-Markers were also included in the desirability assessment, with nearly

equal representation across the Big Five (O:5, C:4, E:5, A:6, ES:5). We obtained correlation coefficients for each participant that indicated the association between WP-SCRs and trait desirability levels for these traits. The mean and median correlations for the sample were both .26, with a considerable range around this value (range: -.41 to .87, sd: .21), replicating our findings in Sample 1 of variation in the tendency to over- or under-claim traits based on desirability. Estimates of this value appear to be reliable, as indicated by a Spearman-Brown adjusted split-half reliability (based on odd- and even-numbered items) of .60; this is lower than was found in Sample 1, as might be expected based on the smaller number of traits available for analysis than in this sample.

Finally, this index of IDR-based bias was associated with existing constructs. High scorers also tended to be high in Self-Esteem ($r = .17, p < .001$) and to score highly on SDR (SDE $r = .21, p < .001$; IM $r = .10, p < .05$).

II.9 Discussion

II.9.1 Desirable Responding: Two Approaches

While a vast literature demonstrates the practical utility of personality as measured by self-report inventories, concerns over the accuracy of such reports are almost as old as the inventories themselves (Edwards, 1957). The range of measures developed to identify the tendency to misrepresent one's personality in such assessments are often discussed as tracking two distinct patterns of such misrepresentation, with some measures (e.g. Impression Management: Paulhus, 1984; Wiggins' Sd scale: Wiggins, 1959) primarily assessing a tendency to deny socially deviant impulses, while others (Self-Deceptive Enhancement: Paulhus, 1984; Edwards' SD scale: Edwards & Walker, 1961) identify those more prone to represent themselves as exceptionally talented and

socially prominent. These patterns of moralistic and egoistic bias are thought to have different motivations (Paulhus & John, 1998) and different eliciting contexts (Paulhus & Trapnell, 2008), showing that causes and patterns of bias in self-report personality are too diverse to be captured by a single construct.

An additional body of intriguing recent work has looked to identify differences in SDR by measuring the extent to which an individual's self-report of personality traits matches the desirability ratings for those traits as provided by a group of raters (Asendorpf & Ostendorf, 1998; Konstabel, Aavik, & Allik, 2006; Saucier, 1994b) rather than relying on purpose-built SDR scales. But whereas these lines of research have identified individual differences in the extent to which people tend to bias self-reports according to universally-applicable patterns of values, a few studies have pursued an alternative path, looking at individual differences in which traits are desired. Sinha and Krueger (1998) showed that differences in perceived trait desirability predicted differences in the self-reported trait levels, though, because their study lacked a comparison to a criterion, it remained unclear whether this finding indicated misrepresentation. More recently, Borkenau et al. (2009) reported that individual differences in the perceived optimal trait level (POTL) were associated with self-reported trait levels. Using peer-reported personality as a criterion measure, this association was shown to partially reflect reality: POTL was significantly associated with the peer-reported level of that trait. At the same time, POTL ratings were more closely matched to self-reported personality than to peer-report. Borkenau et al. (2009) interpreted the latter association as an indication that POTL acts as a source of bias in personality self-report. However, having established POTL as an indicator of true trait variance, it is also

possible that some or even all of the variance found in self- but not peer-report that is shared with POTL was in fact true trait variance. The use of additional indicators of trait levels is needed to clarify the extent to which this relationship between POTL and self-report reflects true variance or bias.

In the present study we expanded on these previous findings. Rather than assessing the effect of perceptions of what trait level was optimal, we looked at the effect of perceptions of how desirable the trait was. This is similar to the procedure used in the development of SDR measures (e.g. Jackson & Messick, 1961), and allowed the use of multiple regression to demonstrate the independence and relative importance of IDR and SDR. As expected, ratings of trait desirability were associated not only with actual trait levels but also with the tendency to misrepresent oneself in self-report of trait levels. Per recommended practice (cf. Paulhus & John, 1998), bias was measured with Self-Criterion Residuals (SCRs). Our observation of an association between SCRs and independent measures of trait levels is consistent with the ability of both self- and peer-report trait levels to incrementally improve each others' validity (Connelly & Ones, 2010). Despite these associations, however, SCRs remained an effective index of bias, particularly with peer-report trait levels as a criterion: a dual-criterion SCR for intelligence was not significantly less associated with trait desirability than was one created using only peer-report. This finding indicates the robustness of IDR. Further, IDR's capacity to predict bias was independent of bias predicted by SDR measures, and for most traits IDR was as an equal or even superior predictor.

Bias in self-reports based on perceived trait desirability was shown to exist as both a mean-level phenomenon (with most participants tending to overclaim traits to the

extent that they view those traits as desirable) as well as an individual difference characteristic, with some participants more prone to this tendency than others. These differences in level of IDR have modest but significant associations with social desirability measures, indicating the potential presence of a common tendency behind the bias of self-reported personality in both a socially and idiographically desirable manner. IDR was also more pronounced in those with more positive global self-evaluations, as indicated by the associations with Rosenberg's Self-Esteem Scale (Rosenberg, 1965). As previous work has indicated that high-scorers on measures of self-worth may be more prone to SDR (e.g. Paulhus & Reid, 1991), our finding that those with high self-worth also overclaim the traits they personally view as desirable is expected.

II.9.2 Comparisons between the two studies

Our results suggest that the effect of individual differences in perceived trait desirability is robust to different methods of assessment. In the first sample participants rated their ideal selves, while participants in Sample 2 were asked to evaluate the desirability of a given trait in other people. To the extent that an individual holds different views of the desirability of a trait based on the individual possessing the trait (themselves versus others), these different assessment methods may partially account for the fact that stronger associations were observed between desirability and personality in Sample 1 than in Sample 2 (although because the samples were demographically quite different as well, the source of this difference remains speculative). Also important for this question are the differences between Sample 1 and 2 in timing of the assessments: the assessments for Sample 1 were contemporaneous, whereas Sample 2 had a two-year gap in between the assessment of personality and trait desirability. Though personality is highly stable

over time (Roberts, Wood, & Caspi, 2008), an interval of this length is likely sufficient to produce some meaningful changes in rank-order personality (Ludtke, Roberts, Trautwein, & Nagy, 2011; Specht, Egloff, & Schmukle, 2011). Any such change would likely attenuate the relationship between desirability and personality, consistent with the smaller associations observed in Sample 2.

Trait desirability ratings thus showed substantially higher associations with both self- and peer personality ratings in Sample 1 than in Sample 2, though with the present data it is not possible to determine whether this derives from the different methods of assessing desirability or instead results from demographic differences between samples or the time gap between assessments in Sample 2. These factors may also have affected the associations between desirability and SCRs. Should individuals differ in the extent to which they believe a trait is desirable in themselves rather than in others, it may be the former that better predicts overclaiming. Similarly, time-discrepancy in the assessment of desirability and of personality would attenuate the connections between the two. Finally, the SCRs in Sample 1 likely contained more true trait variance than did those of Sample 2, as the former sample had only a single peer-report available to identify and remove true trait variance from the SCR. Each of these is consistent with the observation of stronger correlations between desirability and SCRs in Sample 1 than Sample 2. At the same time, it is possible that the SCRs from both studies had lost some variance in self-reports which reflected overclaiming, as those who provided peer-ratings of trait levels tended to be relationship partners, family members, and friends. To the extent that those in such relationships have similar views of which traits are desirable and tend to provide

an overly positive report on their peer, the effects of IDR will be underestimated in studies relying on such reporters.

Another important comparison between the results obtained in our two studies concerns the cross-trait associations. As no such associations were hypothesized and very few of those observed in one sample were replicated in the other we are doubly reluctant to offer substantive interpretation of them. Nevertheless, the reciprocal negative association between the desirability and self-ratings of Openness/Intellect and Conscientiousness (found in both studies) is intriguing: these two traits have opposing associations with political orientation, with liberals characterized by high Openness/Intellect and low Conscientiousness (Hirsh et al., 2010). The observed association is also consistent with Chamorro-Premuzic and Furnham's (2004) compensatory theory, in which those in a competitive work force who are low in cognitive ability may compensate by becoming more Conscientious. As expected, supplementary analyses indicated that those who rank Conscientiousness as highly desirable tended to be lower in cognitive ability, as assessed by our performance measure ($r = -.20, p < .001$).

II.9.3 Considerations for future work

One area for future work concerns clarifying the relations between individual differences in overclaiming based on IDR and scores on measures of SDR and self-esteem. The associations found in Sample 2 are likely an underestimate of both of these associations, due to ceiling effects produced by the limited range of response options available in the personality assessments. Consider the effects of a 5-point Likert-type format in the case of an individual's rating a given trait as extremely desirable or

undesirable. The positive correlation of desirability ratings with self- and peer-reported personality reported above indicate that such individuals will likely be rated extremely on this trait by both themselves and their peers. Such an event was not infrequent: even in Sample 2, which relied on multiple peer raters, roughly 20 percent of our sample had mean-peer ratings of 1 or 5 for any given individual adjective, though aggregated trait ratings such as for a Big Five trait were much less affected by this. At a conceptual level, this can be seen to impair our ability to rate individual differences in the tendency to over- or under-claim trait levels based on desirability, as the traits which individuals were most likely to overclaim were also those that they were not able to, given that their peers assigned them the most extreme score available for that trait.

At an empirical level, the consequences of this limit appear to have had an effect: supplementary analyses on sample 2 treated all traits for which a given participant was given a 1 or 5 mean peer score as missing for that subject when computing WP-SCRs and correlating them with desirability ratings. Although the reduced number of variables available for assessing each subject's tendency to rely on IDR would ordinarily reduce the accuracy of the resulting value for each participant's IDR, in this analysis the resulting index of individual differences in IDR showed slightly stronger associations with the various criterion variables (SDE, IM, and Self-Esteem). Future work on the topic should allow for a wider array of response options, ideally allowing extreme designations such as "top 1%" to allow subjects to identify truly extreme scorers more clearly. This work will make it easier to identify individual differences in IDR, a necessary step towards clarifying the extent of overlap between differences in IDR level and existing constructs. In particular, a more precise estimate of differences in IDR level than was

possible in the present study may allow researchers to differentiate between the motivations and outcomes for different patterns of self-enhancement. For example, individuals who tend to overclaim a trait in proportion to their own views of its desirability might be expected to be high in self-esteem, whereas those who are more prone to overclaim a trait in proportion to consensual views of its desirability might be expected to score high in Impression Management.

Future research may also look to explore the development of the association between perceived desirability and actual trait levels. Supplementary analyses from Sample 2 suggest that some of the typically observed mean-level changes in personality (increases in A, C and ES in adulthood, with decreases in E and O; Roberts et al., 2006) may be matched by changes in the desirability of these traits over time: older participants viewed Extraversion as less desirable (r with age of respondent: $-.14, p < .001$) and Conscientiousness as more desirable ($r = .21, p < .001$), while Agreeableness and Openness/Intellect had insignificant associations in the expected direction ($r = .03, p > .2$; $r = -.02, p > .3$). Contrary to expectations, Emotional Stability was viewed as less desirable by older participants ($r = -.14, p < .001$). Longitudinal data is required to determine whether any associations of this nature might indicate that changes in personality over time are preceded by changes in the perception of the desirability of traits.

Finally, as with any demonstration of the flaws inherent to self-report measures, our study highlights the usefulness of alternate sources of information on trait levels such as peer-report. One unique contribution provided by the present work concerns how certain research questions in personality may derive particular benefit from the use of

non-self-report data to indicate trait levels. Specifically, any studies whose outcomes of interest might reflect value differences between individuals is likely to report systematically exaggerated trait-outcome associations, as those value differences will affect not only the outcome measure but also the self-report of the supposed predictor. For example, a considerable body of literature has developed to characterize the self-reported personality differences of those with different orientations on social and political matters (reviewed in Jost et al., 2003; Sibley & Duckitt, 2008). However, as these domains are intricately connected with values, we should expect that their relationships with personality traits have been systematically distorted due to the effects of IDR, with those at different ends of the spectrum on sociopolitical issues misrepresenting their personality in a fashion consistent with the values reflected in their sociopolitical orientation. Research in this area and others which explore value-related outcomes would thus particularly benefit from the use of non-self-report measures of personality.

II.10 Conclusion

The present study highlights the importance of individual differences in producing biased self-reports. Existing measures effectively capture such differences in the extent of biased responding within certain common patterns (described as egoistic and moralistic bias), but are limited in their ability to assess differences in *which* traits a given individual is more or less prone to misrepresent. Individual differences in perceived trait desirability successfully predicted the traits for which a given individual's self-report would show more or less bias. This process may be more pronounced both among those with high self-esteem as well as those who score high on conventional bias measures. Future work

may look to the role of differences in perceived trait desirability in areas such as personality development and sociopolitically-based personality differences.

Chapter III: Not as different as we want to be:

Attitudinally-consistent trait desirability leads to exaggerated associations between personality and sociopolitical attitudes (Study 2)

Researchers seeking to understand differences in sociopolitical attitudes in terms of associated personality traits have typically relied exclusively on self-report measures of personality. A recently-discovered mechanism of bias in self-reports highlights a particular challenge for this approach. Specifically, there is evidence that individuals tend to report exaggerated levels of a trait to the extent that they view that trait as desirable. In a community sample of 443 participants whose personality was also rated by three peers, differences in sociopolitical attitudes were associated with differences in the extent to which an individual would overclaim a given trait as well as differences in views of the desirability that trait. Further, the tendency to overclaim traits in a manner consistent with one's sociopolitical attitudes was mediated by differences in views of trait desirability. Thus, although meaningful personality differences do exist among those with differing sociopolitical attitudes, those differences may not be as large as people with opposing sociopolitical attitudes might like them to be.

III.1 Introduction

A considerable body of work has stressed the importance of personality in the development of sociopolitical attitudes. Two distinct sociopolitical attitude dimensions have been identified as forming the basis of typical differences between the left and right of the political spectrum (Jost, Nosek, & Gosling, 2008). The first dimension concerns attitudes towards change, contrasting approval of change with a preference for traditionalism and order. This dimension is commonly assessed using Altemeyer's (1988) Right-Wing Authoritarianism (RWA) scale. The second dimension concerns attitudes towards inequality, contrasting egalitarianism with anti-egalitarianism, and is commonly assessed using the Social Dominance Orientation (SDO) scale (Pratto, Sidanius, Stallworth, & Malle, 1994). Conservatives tend to score higher than liberals in both RWA and SDO, indicating that they are more authoritarian and anti-egalitarian. As discussed further below, recent meta-analyses identify the personality traits Openness to Experience, Conscientiousness, Agreeableness, and Intelligence as predictors of RWA and/or SDO (Sibley & Duckitt, 2008; Van Hiel, Onraet, & De Pauw, 2010). In the present study, however, the associations of personality with sociopolitical attitudes were shown to be exaggerated, apparently reflecting a tendency for self-ratings of personality to reflect not only who we are but who we want to be. We found support for a model in which systematic differences in the traits that are desired by people with different sociopolitical attitudes led them to claim to be more different from each other than they actually were.

This bias is present in the research literature because connections between personality traits and attitudes have been established nearly exclusively through the use

of self-reports. Although the utility of self-report assessments of personality is well established (Roberts, Kuncel, Shiner, Caspi, & Goldberg 2007), concerns over the vulnerability of these assessments to distortion have been raised for over half a century (Edwards, 1957). The research described in Study 1 identified a mechanism of bias in self-report that may be of particular concern when identifying the personality bases of sociopolitical orientation: Individual differences in the perception of which traits are desirable leads to bias in personality assessment because people tend to overclaim traits to the extent that they perceive those trait to be desirable. This bias contrasts with “socially desirable responding” (SDR; Paulhus, 2002), in which people tend to claim a particular set of traits that society (rather than the specific individual) finds desirable, and is therefore labeled *idiographically desirable responding* (IDR). To illustrate IDR by example, someone who values Extraversion highly is likely to exaggerate his Extraversion in self-report to a greater extent than someone who does not value Extraversion. Whereas assessment of SDR merely requires knowing what traits are generally desirable, assessment of IDR requires asking each participant about the desirability of each trait.

IDR is relevant to the association of personality characteristics with sociopolitical attitudes because of the well-documented association of sociopolitical attitudes with values, many of which pertain to personality-related characteristics such as politeness and curiosity (Rohan & Zanna, 1996). Whenever people with differing sociopolitical attitudes value such characteristics differently, IDR suggests that they are likely to show differing patterns of misrepresentation in self-reports of personality. For example, Self-Direction (which concerns Openness-related characteristics such as creativity and curiosity) tends

to be valued highly by nonauthoritarians (Rohan & Zanna, 1996), which suggests that nonauthoritarians are likely to overclaim Openness. Thus, correlations between sociopolitical orientation and personality may be driven not only by real differences in personality, but also by the tendency of people with different orientations to exaggerate different traits.

In the present study, a large community sample of adults was used to test the following hypotheses: (1) sociopolitical attitudes are associated with overclaiming in personality assessments, (2) when sociopolitical attitudes predict overclaiming for a trait, they also predict views of the desirability of that trait, (3) the association of sociopolitical attitudes with overclaiming is mediated by self-ratings of trait desirability. We focus on traits established by meta-analytic work (Sibley & Duckitt, 2008; Van Hiel et al., 2010) as relevant to social attitudes: Agreeableness, Conscientiousness, Openness, and Intelligence.

III.2 Method

III.2.1 Participants

Participants were members of the Eugene-Springfield Community Sample, a predominantly middle-aged group ($M = 51$ years, $SD = 12.36$, range = 18–80) drawn from lists of local homeowners. Participants were predominantly (97%) Caucasian and reported a range of educational attainment, with a median of two years of post-secondary education. Surveys were completed by mail over 14 years in exchange for money, beginning in 1994. We used all participants for whom data were available for self-reported personality, trait desirability, and three peer-reports of personality. After excluding two participants whose responses indicated obvious inattention (all 97 trait

desirability ratings were marked as “neither desirable nor undesirable), 443 participants were included. For the analysis of intelligence, 383 participants who completed all relevant measures were included.

III.2.2 Measures

III.2.2.1 Big Five

Big Five personality traits were assessed with the 44-item Big Five Inventory (BFI; John & Srivastava, 1999) and Saucier’s (1994a) 40-item Mini-Markers (MM). These measures were also completed independently by three peers who knew the participants well. All instruments utilized 5-point Likert scales. Scores for Big Five personality for both self- and peer-report were obtained by taking the means of all items for each trait from both the BFI and MM, yielding alphas between .84 and .94. As described in Study 1, a subset of Openness items related to intelligence (e.g., “insightful,” “unintelligent”) was selected to represent self-reported intelligence levels.

III.2.2.2 Desirability

Participants were asked to rate “how desirable or undesirable you feel it is for others to be or act this way” for a list of 97 characteristics using a 9-point Likert scale, with responses translated into trait desirability measures for the Big Five based on Saucier and Goldberg’s (1996) analysis of which of these adjectives fell within each Big Five construct. To compute the desirability of intelligence, a subset of Openness items was selected as described in Study 1. As described in Study 1, scores were adjusted to remove the effect of response sets such as extreme responding (in which some people tend to use the extreme ends of the scale and others the middle responses), which would artificially make some participants appear to value all traits more than other participants. Partialling

out the other four Big Five traits from each score ensured that scores reflected the unique desirability of each trait for each participant. (Because ratings of intelligence were assessed with a subset of Openness items, only the four remaining Big Five traits were partialled from intelligence.) Supplementary analyses using raw (unpartialled) desirability ratings yielded substantively the same pattern of associations that is reported below.

III.2.2.3 Intelligence

383 participants completed Cattell's 16 Personality Factor Questionnaire (16PF; Conn & Rieke, 1994). Factor B of this instrument is a 15-item intelligence test including knowledge and reasoning problems with objectively correct answers in a multiple-choice format. Previous research has shown that scores correlate well with Full Scale IQ from the Wechsler Adult Intelligence Scale – Revised, $r = .57$ (Abel & Brown, 1998).

III.2.2.4 Social attitudes

Participants completed a 14-item version of Altemeyer's (1988) RWA scale and a 10-item version of Pratto et al.'s (1994) SDO scale. RWA assesses a preference for convention, submission to established authorities, and willingness to punish social deviants, as exemplified by the item "Our country will be great if we honor the ways of our forefathers, do what the authorities tell us to do, and get rid of the 'rotten apples' who are ruining everything." SDO assesses an individual's willingness to accept group-based inequality, and is assessed with items like "It's OK if some groups have more of a chance in life than others."

III.2.3 Analysis

Previous work on identifying social and idiographic influences on overclaiming has relied on self-criterion residuals (SCRs) to indicate bias in self-reports (Paulhus & John, 1998; Study 1 of this dissertation). SCRs are obtained by predicting self-reported trait levels using one or more criteria, with the residual variance not predicted by the criteria representing bias. Though most previous work relied exclusively on peer-report as a criterion, we used objective performance for the one trait (intelligence) for which such data were available.

To test our hypothesis that sociopolitical attitudes predict SCRs because differences in attitudes are associated with differences in views of the trait desirability, we employed structural equation modeling as shown in Figure 1. For Openness, Conscientiousness, and Agreeableness, each peer-report served as a criterion. For intelligence we created three indicators, each comprising five Factor B items, by summing scores for every third item. The three variables on the right of the figure constitute a standard mediation analysis, and the inclusion of the latent variable allowed us to separate variance due to bias from true trait variance. True trait variance is indicated by the agreement among peer and self-ratings and is likely to affect all variables in the model, including trait desirability because (as shown in Study 1) people tend to value traits that they actually possess. Because most of the variance in self-reports that is due to true trait levels (i.e. that is shared with the other indicators of the trait) is accounted for by the path from the latent True Trait variable to Self, any other paths leading to Self are predicting variance that is roughly equivalent to the SCRs represented in Study 1 and in Table 6 of this study. (However, because the model additionally incorporates Sociopolitical Attitudes and Trait Desirability as indicators of True Trait Levels, the

variance remaining in Self in this model will not be identical to that remaining in the SCRs). Because a considerable portion of the reliable variance in self-reports is removed from both SCRs and the Self indicator in the model, the proportion of variance remaining in each that is due to measurement error will be relatively high, attenuating any potential relationships of both SCRs and the Self indicator with other constructs of interest. Finally, the number in parentheses is the value for the path from Sociopolitical Attitudes to Self when the mediator, Trait Desirability, was not included in the model.

III.3 Results

Table 6 shows correlations between the constructs of interest. Matching the results of previous meta-analyses, in self-reports authoritarians were high in Conscientiousness but low in Openness and intelligence, and anti-egalitarians were low in Agreeableness and Openness (Sibley & Duckitt, 2008; Van Hiel et al., 2010). Very similar results were observed for peer reports and objective performance, although RWA was not significantly correlated with peer-reported Conscientiousness.

Insert Table 6 about here

We observed comparable associations between sociopolitical attitudes and personality as indicated by self- and peer-report as well as objective performance, indicating some degree of validity for associations between personality and sociopolitical attitudes. Nonetheless, we also found evidence supporting our first hypothesis, that these associations were overstated in self-reports. Table 6 includes correlations with SCRs

indicating that RWA was associated with overclaiming Conscientiousness but underclaiming Openness and intelligence. SDO predicted underclaiming Agreeableness.

Consistent with our second hypothesis, sociopolitical orientation predicted trait desirability for those traits where sociopolitical views predicted overclaiming. Authoritarians tended to view Conscientiousness as particularly desirable while placing little value on Openness and intelligence. Anti-egalitarians tended to place a low value on Agreeableness, with more modest associations observed between SDO and views of the desirability of Openness, intelligence, and Conscientiousness.

Our third hypothesis was tested using the mediation model shown in Figure 1. For each instance in which a sociopolitical attitude predicted overclaiming of a trait, we tested whether the effect was transmitted through differences in perceived trait desirability. Table 2 presents standardized path coefficients and fit indices from these models. RMSEA and CFI values indicated the model fit well in all cases. Mediation was tested using the bootstrap method (Shrout & Bolger, 2002), and Trait Desirability mediated the effect of Sociopolitical Attitudes on Self for all cases at $p < .05$. That is, the association between sociopolitical attitudes and misrepresentation on certain traits could be explained by their associations with perceptions of the value of those traits. Sociopolitical attitudes had no significant direct effect on self-ratings, after accounting for their association with desirability, except in the case of RWA and intelligence, for which mediation was only partial.

Insert Table 7 and Figure 1 about here

Finally, in addition to predicting bias, sociopolitical attitudes were also associated with true differences in personality, as indicated by significant paths from the latent True Trait to Sociopolitical Attitudes in every model except for the one including Conscientiousness and RWA.

III.4 Discussion

A large and growing research literature highlights the connections between self-reported personality measures and sociopolitical attitudes, with recent work looking to these connections to explain changes in attitudes over time (Sibley & Duckitt, 2010) and even exploring common genetic roots between personality and attitudes (Verhulst, Hatemi, & Martin, 2010). Our results suggested that the connection between personality and sociopolitical attitudes can contain elements of both truth and fiction. Although attitudes did show connections to true latent trait levels in a way that closely matched results from recent meta-analyses of self-reports, attitudes were also associated with the tendency to over- or under-claim many of these traits. Further, sociopolitical attitudes were correlated with perceptions of how desirable certain traits were. Finally, mediation analysis showed that these associations between sociopolitical attitudes and exaggeration of personality traits could be explained by differences in perceptions of trait desirability that were associated with sociopolitical attitudes. This suggested that these attitudes led to distortions in self-rated personality traits because they were associated with perceptions of the desirability of those personality traits.

These results suggest that people with differing views on social and political issues may not be as different in personality as previously thought – or as different as some of those with opposing views might like to believe. For example, the elevated Conscientiousness that is taken to characterize authoritarians (Sibley & Duckitt, 2008) appeared only in self-report in our study, with the high value authoritarians ascribe to Conscientiousness apparently inspiring them to overclaim the trait. Similarly, those low in SDO (egalitarians) overclaimed their Agreeableness apparently as a result of the high value they placed on the trait, and nonauthoritarians did the same for Openness. Importantly, this pattern of results was also detected in a trait for which an objective assessment was performed, reducing any concern that our results are merely artifacts of inaccurate peer-ratings: nonauthoritarians tended to claim higher levels of intelligence than were indicated by objective performance, and this effect was mediated by the high value that nonauthoritarians placed on intelligence.

The present results additionally highlight just how far apart those with different sociopolitical views are when it comes to the traits they view as desirable. Future work may explore the various effects these differences might have on the political process, including candidate selection. For example, whereas left- and right-wing voters tend to differ modestly in IQ and Openness (Jost et al., 2003), recent Democratic presidents have been markedly higher on these traits than recent Republican presidents (Simonton, 2006), perhaps reflecting the high value placed on these traits by the left. Similarly, while center-right voters in Italy tend to score modestly higher than center-left voters on Conscientiousness, this gap appears more pronounced among Italian political leaders

(Caprara, Barbaranelli, Consiglio, Picconi, & Zimbardo, 2003; Caprara, Barbaranelli, & Zimbardo, 1999).

More generally, our results provide a concrete demonstration of the need for caution in interpreting any reported connections between self-reports of personality and variables, like sociopolitical attitudes, that may reflect the values of individuals. Differences in the values placed on different traits have the power to bias self-reports of personality such that they show an exaggerated association with other value-related constructs. Future research looking to connect personality to value-laden constructs would benefit from the use of alternative measures of personality such as peer reports (Connelly & Ones, 2010) and objective measures like intelligence tests.

Chapter IV: Authoritarianism as a Personality Trait: Evidence from a Longitudinal Behavior Genetic Study (Study 3)

Authoritarianism has long been conceived of as a highly stable personality trait (Adorno et al., 1950; Altemeyer, 1981), though recent accounts have argued that authoritarianism is too malleable to justify this conception. We provided a test of the trait conception of authoritarianism by measuring its stability in a community sample of twins over a 15 year period, and by identifying the source of any stability with biometric modeling. Our results showed that authoritarianism exhibited a high degree of rank-order stability ($r = .74$). Biometric analyses indicated that this stability derived primarily from genetic influences, with changes in authoritarianism due to the unique experiences of the individual. In both of these respects, our results were highly comparable to those reported for other personality traits in previous work, indicating support for the trait conception of authoritarianism. Other results of note included a higher degree of stability among the more educated portion of the sample, supporting a hypothesis by Krosnick and Alwin (1989).

IV.1 Introduction

Authoritarianism was represented as a highly stable personality trait in both the original conception put forward by Adorno and colleagues (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950) as well as the widely-used revision of the construct offered by Altemeyer (1981; 1988; 1996). Recent approaches have challenged this view, with

several authors (e.g. Sibley, Wilson, & Duckitt, 2007; Van Hiel, Pandelaere, & Duriez, 2004) arguing that authoritarianism is too susceptible to change to be considered a personality trait, and is instead better characterized as a characteristic adaptation (McAdams & Pals, 2006) or surface trait (Asendorpf & Aken, 2003). A number of conceptualizations of personality traits have been offered in recent years, providing some guidelines regarding the features that a psychological characteristic should have in order to be considered a “trait.” For example, many theorists suggest that traits must exhibit a high degree of stability over time (e.g. Asendorpf & Aken, 2003; McAdams & Pals, 2006; McCrae & Costa, 2008). Most recent challenges to authoritarianism’s status as a trait have focused on its supposed instability, though as longitudinal data on the topic has been scarce and dominated by student samples there is a clear need for additional data from community based samples. In the present study we analyzed data from a community-based longitudinal twin study of adults, seeking to ascertain both the stability of authoritarianism over a 15 year interval and the extent to which genetic and environmental sources contribute to that stability.

IV.1.1 Previous research on the stability of authoritarianism and other personality traits

In one study frequently cited among challenges to the trait conception of authoritarianism, individuals asked to anticipate their beliefs during an apocalyptic future imagined their future selves to be moderately more authoritarian than did those asked to anticipate their beliefs in a future with more stable economic and social conditions (Duckitt & Fisher, 2003). Other studies went beyond the hypothetical, finding that modest-to-moderate changes on an abbreviated authoritarianism measure were successfully predicted by authoritarianism-related constructs such as Openness to

Experience (Sibley & Duckitt, 2010) and perceptions of the world as a dangerous place (Sibley et al., 2007). The interpretation of results from these findings has occurred within a problematic conception of personality traits, in which traits are suggested to be “immutable” (Sibley et al., p. 358) and “invariant across situations” (Sibley et al., p. 367). Both of these conditions are more stringent than is required by contemporary theoretical accounts of traits cited above, and as discussed below, such conditions would exclude from the trait domain not only authoritarianism but also constructs such as extraversion and conscientiousness.

The claims regarding authoritarianism’s purportedly excessive mutability are not typically accompanied by an acknowledgment of previous work highlighting the construct’s longitudinal stability, or by an effort to place the experimental results on authoritarianism within the context of the broader literature on personality stability. Previous work on the former topic includes Altemeyer’s (1996) report that a group of alumni assessed first as freshman and then again 18 years later maintained an impressive degree of rank-order stability over time, with a correlation of .59 between the two assessments. Altemeyer (1996) reports scores that have not been standardized, and he does not report the variance for each assessment. However, based on the variance observed in the same instrument in different samples, mean-level changes appear to have been modest, with authoritarianism scores of participants exhibiting an average decline of approximately a third of a standard deviation over this period. Altemeyer (1996) also reported very similar results on a different sample assessed 12 years after their freshman assessment, reporting a test-rest correlation of .62. When put in the context of recent meta-analyses regarding the stability of personality, authoritarianism appears to be more

stable over time than other personality traits, not less. For example, the correlation between trait levels at ages 18 and 22 is only expected to be .54, with even more modest correlations expected for longer time intervals (Roberts & DelVecchio, 2000). Similarly, Roberts, Walton, & Viechbauer (2006) report that increases of half a standard deviation or greater are expected for Openness, Conscientiousness, Emotional Stability, and Social Dominance during the ages in which Altemeyer's (1996) participants showed a mean shift of a third of a standard deviation in authoritarianism.

Nevertheless, there are important limitations to the generalizability of Altemeyer's (1996) results concerning the stability of authoritarianism over time. His studies considered change during only one segment of the life course (early adulthood), and relied on exclusively on college-student populations. College-educated individuals differ from those without post-secondary degrees in their levels of authoritarianism (McCourt et al., 1999), and Krosnick and Alwin (1989) hypothesized that advanced education may act to solidify political attitudes, which would lead to elevated levels of stability among educated samples. Studies using samples with more diverse educational backgrounds are clearly needed.

Because the rank-order stability for personality traits typically increases with age (Roberts & DelVecchio, 2000), a conception of authoritarianism as a personality trait leads to the prediction that samples older than that used by Altemeyer should show greater levels of stability. This stability may be especially pronounced among more educated individuals, as hypothesized by Krosnick and Alwin (1989). In addition, while authoritarianism is known to be substantially influenced by genetic factors (McCourt et al., 1999), we are aware of no study which has looked at the role of genetic and

environmental influences on stability and change in authoritarianism over time. Previous work on personality traits (e.g. Johnson, McGue, & Krueger, 2005) has shown that genetic factors contribute almost exclusively to rank-order stability in traits over time, while nonshared environmental factors account for rank-order change.

The data used in the present study were collected from a community-based sample of twins who were assessed for authoritarianism in middle-age and then again 15 years later. Based on the above review, we derived the following three expectations from a conception of authoritarianism as a personality trait:

- 1) Due to the age of this sample, authoritarianism should show higher levels of rank-order stability than it did in younger samples;
- 2) This stability should be primarily genetic in origin;
- 3) Genetic influences will contribute primarily to stability and not to change, while nonshared environmental influences will contribute both to stability and to change

To the extent that these expectations are met, the conception of authoritarianism as a personality trait will have been supported.

IV.2 Methods

IV.2.1 Participants and Measures

Participants were members of the Minnesota Twin Registry (MTR), a birth-record based registry of twin pairs born in Minnesota described at length by Krueger and Johnson (2002). The present study assessed those members of the registry who were

members of same-sex twin pairs born between 1947 and 1955, who completed two assessments of interest.

IV.2.1.1 Assessment 1

As described in greater detail by McCourt et al. (1999), 2,800 MTR participants completed the 1986 version of the RWA scale (Altemeyer, 1988) between 1990 and 1993. Participants used a nine-category Likert response format to provide responses to 30 items assessing the three facets of the authoritarianism construct put forward by Altemeyer: conventionalism, authoritarian submission, and authoritarian aggression. An example item representing all three of these facets is “Our country will be great if we honor the ways of our forefathers, do what the authorities tell us to do, and get rid of the ‘rotten apples’ who are ruining everything.” The alpha reliability of the measure was .94. At this time, participants also indicated the years of education they had achieved. Of the participants included in the present analysis (those who completed both assessment 1 and assessment 2), the median participant had completed two years of post-secondary education (Mean = 14.31, *SD* = 2.31, range 10 – 21).

IV.2.1.2. Assessment 2

A comprehensive assessment of political attitudes was performed in 2008 and 2009 on MTR participants born between 1947 and 1955, as described in detail by Funk et al. (2012). The assessment included Zariksson’s (2005) abbreviated form of Altemeyer’s RWA measure, which assesses 15 items with a seven-category Likert response format. Though these items are thematically highly similar to Altemeyer’s (1988) measure, no items were exactly identical, preventing a meaningful comparison of mean-levels

between the two assessments. 1,327 participants completed this assessment, including 540 (53% female) who completed Assessment 1. This included 131 MZ twin pairs and 86 DZ twin pairs, where members of an additional 45 MZ and 49 DZ twin pairs provided complete information at one assessment with only one member of the pair completed the other assessment. The alpha reliability of this measure was .87.

IV.2.2 Analysis

Twin models make use of the differences in the genetic similarities between MZ and DZ twin pairs to quantify the relative contributions of genetic and environmental factors to a given phenotype. This typically involves decomposing phenotypic variance into variance due to additive genetic effects (a^2 : the summed contribution of genes across loci) as well as shared (c^2) and nonshared (e^2) environmental effects. Shared environmental effects produce similarity between twin pairs regardless of zygosity, while nonshared environmental effects produce uniqueness among members of a twin pair. Measurement error and state fluctuations are also represented as nonshared environmental effects.

We used a Cholesky model to estimate the biometric contributions to RWA at each time point as well as the extent to which these contributions are consistent over time. A simplified version of this model is presented in Figure 1. This model also allowed us to estimate the extent to which genetic and environmental variance at Time 2 is shared with Time 1 (indicating stable influence over time), or is instead unique to Time 2 (indicating it is responsible for change). We fit our models to the raw data using Full Information Maximum Likelihood, which provides efficient and consistent estimates in

the presence of missing data (Little & Rubin, 1987). The assumptions required for the use of this model on twin data are discussed more fully in Johnson (2007). One issue of particular importance concerns the presence of assortative mating, which can provide a downward bias to heritability estimates if not accounted for in the model. Previous work on this sample (McCourt et al., 1999) and others (Martin et al., 1986) has demonstrated the importance of assortative mating for sociopolitical attitudes. Based on the results of McCourt et al. (1999), we incorporated an assortative mating coefficient of .4 into our model.

IV.3 Results

IV.3.1 Phenotypic stability

Phenotypic stability for RWA was very high, with Time 1 and Time 2 scores correlating .74 (95% confidence interval: .71 - .77). This stability was particularly pronounced among the more educated segment of the sample. Among those with 14 or more years of education (N=285), the correlation between Time 1 and Time 2 scores was .78, significantly higher than the correlation of .64 among those with 13 or fewer years of education (N=240; $p < .001$). The more educated portion of the sample tended to exhibit more consistent responses within a given assessment: alphas were higher for Time 1 and Time 2 for the more educated sample (.95 and .90, respectively) than for the less educated sample (.90 and .80, respectively). However, these differences in reliability could not account for the greater stability: after correcting for attenuation due to measurement error, the correlation between Time 1 and Time 2 scores was still

significantly ($p < .01$) higher among the more educated portion of the sample ($r = .84$) than among the less educated ($r = .76$).

The correlation between authoritarianism at Time 2 and years of education was $-.38$ ($p < .001$), consistent both with previously reported results using the Time 1 assessment of this sample as well as results from a recent meta-analysis (McCourt et al., 1999; Van Hiel, Onraet, & De Pauw, 2010). An individual selecting the neutral (central) response for each question from the Time 2 assessment would obtain a score of 60; the mean score among the less educated group (64.26; $sd = 14.20$) indicated a significant tendency to provide authoritarian-leaning responses ($p < .001$), while the more educated group (53.72, $sd = 17.87$) tended to endorse slightly nonauthoritarian views ($p < .001$), though the variance was significantly greater in the more educated sample ($p < .001$).

IV.3.2 Biometric results

Results from Table 1 indicate that MZ twin pairs were more similar to each other than were members of DZ twin pairs, pointing to the importance of genetic factors for authoritarianism. Cross-twin, cross-time correlations were obtained by correlating the Time 1 scores of one member of a twin pair with the Time 2 scores of the other member of the pair. MZ twins are more similar to each other over time than are DZ twins, indicating the presence of genetic influence on the stability of RWA over time.

Results from the Cholesky model (presented in Table 2) provide a comparable pattern of results as that derived from the twin correlations. Additive genetic variance accounted for roughly half of the variance at each assessment, with nonshared environment accounting for most of the remainder. Consistent with research on other

personality traits (e.g. Johnson et al., 2005), genetic factors accounted for 58% of the stability over time in RWA, with lesser roles played by the shared and nonshared environment. Finally, the genetic contributions to Time 1 RWA overlapped completely with genetic contributions to Time 2 RWA, indicating that the modest amount of rank-order change in RWA during this time period was not due to genetic factors. A similar result was observed for shared environmental effects. In contrast, nonshared environmental contributions to Time 1 variance overlap only moderately (.36) with contributions to Time 2 variance.

IV.4 Discussion

Our results were consistent with the conception of authoritarianism as a highly stable personality trait. While our use of different measures of authoritarianism at the two assessments might be expected to attenuate the correlation between them, we nevertheless observed a very high degree of rank-order stability, with a correlation of .74 over a 15 year interval. This is identical to the expected 10-year stability of .74 for personality traits for adults in their 50s reported by Roberts et al. (2006). This stability was significantly greater among the more educated members of our sample, supporting Krosnick and Alwin's (1989) hypothesis that advanced education leads to a solidification of attitudes. As Altemeyer's (1996) work on the long-term stability of authoritarianism had relied exclusively on student samples, it may provide overestimates of the stability of the trait in the population at large during young adulthood.

This study also provided the first genetically-informative longitudinal study of authoritarianism. We found that rank-order changes in authoritarianism derived from

nonshared environmental factors, while genetic influences were the primary contributor to rank-order stability in the phenotype. That is, phenotypic stability was primarily due to genetic factors ($a^2 = .58$). These results are consistent with those reported in a study of a portion of the present sample in which personality was assessed twice over a five-year interval with the Multidimensional Personality Questionnaire (MPQ; Tellegen & Waller, 2008): Johnson et al. (2005) reported that genetic factors contributed nearly exclusively to stability and not change, with nonshared environment contributing to both stability and change. Thus, while both authoritarianism and the MPQ traits have been observed to be highly stable over time, with genetics accounting for much of that stability, the changes that did occur in scores for these measures reflected both measurement error and the unique experiences of the individual.

Longitudinal twin studies using the MPQ are particularly useful for evaluating how authoritarianism's stability and biometric basis compare to that of other personality traits, as the MPQ Traditionalism scale is highly comparable to Altemeyer's authoritarianism measure (Ludeke, Johnson, & Bouchard, 2013). Consistent with the conceptualization of authoritarianism and related characteristics as personality traits, two recent studies showed that the phenotypic stability of Traditionalism was similar to that observed for other MPQ traits (Blonigen et al., 2008; Johnson et al., 2005). At the same time, the unique environment may be less important for authoritarianism and related traits than for other personality traits: twin correlations for authoritarianism in this study and for the Traditionalism scale in Johnson et al. (2005) were at the high end of what has been reported for personality traits (reviewed by Johnson, Vernon, & Feiler, 2008), and estimates for the unique environment were significantly smaller for Traditionalism than

for most other MPQ traits in Blonigen et al. (2008). This result may be substantive rather than simply reflecting the role of measurement error, given that the reliability for Traditionalism is comparable to other MPQ scales (whether assessed with Cronbach's Alpha or with 30-day test-retest correlations; Johnson et al., 2005). Rather than being a mutable characteristic that is highly responsive to the experiences of the individual, then, authoritarianism and related traits may be particularly effective exemplars of stable traits unaffected by the individual's unique experiences.

Efforts to resolve the current dispute over the nature of authoritarianism require attention not only to results such as those presented here, but also an eye towards contemporary conceptions of personality traits.

In particular, a recognition of the susceptibility of personality traits to environmental influence (Roberts et al., 2006) indicates that the moderate mean-level changes indicated by other studies of authoritarianism are not inconsistent with a trait conception of the construct. For example, the observation that moderate decreases in traits such as negative emotionality accompany occupational success has not prompted any dispute over negative emotionality's status as a personality trait (Roberts, Caspi, & Moffitt, 2003). Similarly, if the participants in Duckitt and Fisher's (2003) study were correct in hypothesizing that their authoritarianism scores would moderately increase should society descend into chaos, this is best seen as yet another example of how an individual's personality traits may undergo predictable shifts in response to changes in his environment – not as a challenge to authoritarianism's status as a trait. At the same time, the demonstration that authoritarianism exhibits a high degree of genetically-influenced stability does not trivialize those changes in authoritarianism that do occur,

particularly given the societal implications of such changes. For example, citizens of European states hit particularly hard by the recent global economic crisis have shown increased sympathy for authoritarian governments (Berglof, 2011), where this shift has been accompanied by corresponding election results and political movements in many of the affected nations. Thus, while authoritarianism may be no more malleable than any other personality trait, the identification of contributors to change in the trait remains a particularly urgent project deserving of continued attention in future research.

Future research is also needed to address other arguments against the trait conception of authoritarianism. In particular, the extent to which the effects of authoritarianism on behavior are “invariant across situations” (Sibley et al., 2007; p. 367) should be compared to the extent to which the effects of traits like extraversion on behavior may be affected by situational context. With regard to authoritarianism, although it is generally linked to greater rates of intergroup bias (Altemeyer, 1988; 1996; Meeus, Duriez, Vanbeselare, Phalet, & Kuppens, 2009), this effect can be intensified when individuals have recently considered a range of threatening scenarios, including their death, property loss to burglary, and the prospect of social isolation (Navarrete, Kurzban, Fessler, & Kirkpatrick, 2004). That is, the effect of authoritarianism on outcomes of interest may be exaggerated or muted based on other factors.

Of course, moderators of the effects of conventional personality traits have long been known. For example, the effects of extraversion can be suppressed or exaggerated by the nature of a social interaction: Monson and colleagues (1982) reported that when participants were either actively solicited to socialize or actively discouraged from socializing, extraversion was not predictive of talkativeness, whereas in a neutral

condition lacking both solicitation and discouragement, extraversion was significantly predictive of talkativeness. There is thus nothing about the demonstration that a given characteristic has non-uniform expression across different situations that threatens its status as a personality trait. Indeed, as Tellegen and Waller (2008) note, traits “are emphatically not construed as generalized ‘situation-free’ action tendencies, but as tendencies to behave in certain ways in certain situations” (p. 262).

Rather than expecting, as Sibley et al. (2007) appear to, that personality has wholly unconditional effects on behavior, we might expect that the trait in question simply leads to a consistent direction of effect given a particular class of situations. This criterion would still be quite stringent – for example, it would require that no context should be sufficient to cause extraverts to socialize less than introverts without challenging the meaningfulness of “extraversion” as a trait. In the case of authoritarianism, its effects could be argued to be unconditional if authoritarians show consistently greater tendencies to submit to authorities they recognize as legitimate. To sustain a challenge to authoritarianism’s status as a personality trait on the grounds of the purported context-dependent nature of its effects on behavior, researchers might explore whether there are any contexts that make authoritarians more willing than nonauthoritarians to rebel against authorities perceived to be legitimate.

Table 1. *Associations of trait desirability with personality and self-criterion residuals in student sample*

		Trait Desirability				
		O	C	E	A	ES
Self-Report	O	.64	-.14	.00	.03	-.11
	C	-.12	.65	.03	.03	-.15
	E	-.09	-.15	.62	.05	-.14
	A	-.01	.00	-.01	.66	-.07
	ES	-.06	-.15	.03	-.04	.44
Peer-Report	O	.32	-.05	-.07	.12	.01
	C	-.06	.36	.01	.03	-.06
	E	-.03	.02	.27	-.08	.02
	A	.00	-.02	.00	.27	.12
	ES	-.03	-.08	.00	-.08	.33
Self-Criterion Residuals	O	.55	-.13	.04	-.03	-.12
	C	-.11	.55	.03	.02	-.13
	E	-.08	-.17	.56	.09	-.17
	A	.00	.01	-.01	.60	-.12
	ES	-.06	-.12	.04	.00	.33

Note: Each value is a correlation across 176 targets of their ratings of the desirability of a trait with trait-levels as indicated by self- and by peer-report as well

as with self-criterion residuals for a trait. Predicted relationships are in bold, and are all significant at $p < .001$. Unpredicted relationships significant at $p < .05$ (uncorrected for multiple testing) are in italics. O = Openness/Intellect, C = Conscientiousness, E = Extraversion, A = Agreeableness, ES = Emotional Stability.

Table 2. *Associations of trait desirability with trait levels and self-criterion residuals in adult community sample*

		Trait Desirability					
		O	C	E	A	ES	Attract
Self-Report	O	.51	-.10	-.03	-.05	.00	.01
	C	-.09	.34	-.02	.03	-.04	-.05
	E	-.05	.00	.27	-.01	-.01	-.08
	A	-.04	-.05	.07	.40	-.14	-.07
	ES	.04	.01	.01	-.09	.22	.05
	Attract	.01	-.08	.15	-.02	.10	.13
	Peer-Report	O	.40	-.18	-.04	-.02	.07
C	-.10	.21	-.05	.05	-.03	-.04	
E	-.08	-.04	.27	.00	.01	-.09	
A	-.02	-.08	.03	.22	-.04	.02	
ES	.01	.04	-.01	-.09	.09	.09	
Attract	.00	-.06	.03	.18	.03	.06	
Self-Criterion Residuals	O	.35	.00	-.01	-.05	-.05	.02
	C	-.05	.28	.01	.01	-.03	-.03
	E	.00	.03	.13	-.01	-.03	-.03
	A	-.04	-.02	.06	.34	-.14	-.08
	ES	.04	-.01	.02	-.06	.20	.01
	Attract	.01	-.06	.14	-.08	.09	.11

Note: Each value is a correlation across 541 targets of their ratings of

the desirability of a trait with trait-levels as indicated by self- and by peer-report as well as with self-criterion residuals for a trait. O = Openness/Intellect, C = Conscientiousness, E = Extraversion, A = Agreeableness, ES = Emotional Stability, Attract = Physical Attractiveness. Predicted relationships are in bold. All associations of .13 or greater are significant at .001; all associations of .08 or greater are significant at $p < .05$ (uncorrected for multiple comparisons).

Table 3. *Intelligence: Associations Between Desirability, Trait Levels, and Self-Criterion Residuals (SCRs)*

	Desirability	Self-rating	Peer-rating
Self-rating	0.48		
Peer-rating	0.38	0.53	
Cattell	0.41	0.34	0.39
SCR - Peer	0.33	X	X
SCR - Cattell	0.35	X	0.42
SCR - Peer & Cattell	0.27	X	X

Note: Cells are marked with an X if one variable was used to compute the other. All values are significant at $p < .001$.
Cattell = Cattell's Factor B.

Table 4. *Correlations of the BIDR scales with self-criterion residuals*

		Impression	Self-Deceptive
		Management	Enhancement
	Openness/Intellect	-.02	.13
	Conscientiousness	.16	.32
Self-	Extraversion	.02	<i>.09</i>
Criterion	Agreeableness	.28	.05
Residuals	Emotional Stability	.30	.39
	Attractiveness	-.07	.07
	IQ	.00	<i>.12</i>

Note: Associations in bold are significant at $p < .01$, associations in italics are significant at $p < .05$ (uncorrected for multiple testing).

Table 5. *Overclaiming Predicted in Multiple Regression Using BIDR Scales and Trait Desirability*

		Beta	Sig
	IM	-.06	.168
O	SDE	.15	.001
	Desirability	.37	.000
	IM	.00	.963
C	SDE	.29	.000
	Desirability	.24	.000
	IM	.00	.934
E	SDE	.08	.084
	Desirability	.12	.007
	IM	.24	.000
A	SDE	.00	.913
	Desirability	.31	.000
	IM	.20	.000
ES	SDE	.29	.000
	Desirability	.17	.000
	IM	-.10	.044
Attract	SDE	.11	.017
	Desirability	.09	.037
	IM	-.07	.183

IQ	SDE	.14	.004
	Desirability	.26	.000

Note: Participant scores for self-criterion residuals for each trait were predicted simultaneously by the BIDR scales and the participant's rating of the desirability of this trait. IM = Impression Management, SDE = Self-Deceptive Enhancement, O = Openness/Intellect, C = Conscientiousness, E = Extraversion, A = Agreeableness, ES = Emotional Stability, Attract = Physical Attractiveness.

Table 6. *Correlations of sociopolitical attitudes with personality, self-criterion residuals, and trait desirability ratings*

		RWA	SDO
Self-ratings	A	.01	-.30
	C	.12	-.05
	O	-.38	-.10
	IQ	-.35	-.06
Peer-ratings	A	.02	-.19
	C	.06	-.03
	O	-.35	-.19
Objective	IQ	-.36	-.03
Self-Criterion Residuals	A	.00	-.23
	C	.10	-.04
	O	-.22	.01
	IQ	-.24	-.07
Desirability	A	.06	-.39
	C	.33	.20
	O	-.47	-.21
	IQ	-.42	-.18

Note: Values in bold are significant at $p < .05$ (without adjustment for multiple testing). A = Agreeableness, C = Conscientiousness, O =

Openness, IQ = Intelligence.

Table 7. Standardized path weights and fit indices from the model in Figure 1

		True → S.A.	True → Des	S.A. → Des	True → Self	True → Crit. 1	True → Crit. 2	True → Crit. 3	S.A. → Self	Des → Self	RMSEA	CFI
C		.07	.18*	.32**	.54**	.66**	.67**	.61**	.00 (.09*)	.24**	.000	1.000
RWA	O	-.42**	.31**	-.34**	.58*	.60**	.69**	.66**	-.02 (-.11*)	.27**	.041	.993
	IQ	-.43**	.35**	-.27**	.19**	.71**	.76**	.61**	-.13* (-.22**)	.33**	.000	1.000
SDO	A	-.25**	.19**	-.34*	.52**	.61**	.53**	.57**	-.08 (-.16**)	.22**	.028	.994

Note: In the models containing C, O, and A, criterion values are provided by peer reports. In the model containing IQ, criterion values are provided by 5 item clusters from Factor B. The value in parentheses represents the path coefficient in the reduced model, which omits Trait Desirability. S.A. = Social Attitudes, Des = Desirability, Crit. = Criterion, A = Agreeableness, C = Conscientiousness, O = Openness, IQ = Intelligence. ** = $p < .01$; * = $p < .05$

Table 8. *Intraclass correlations in authoritarianism scores between members of twin pairs*

	MZ			DZ		
	Time 1	Time 2	CTCT	Time 1	Time 2	CTCT
<i>r</i>	.67	.59	.63	.49	.43	.43
	(.62, .72)	(.52, .65)	(.57, .68)	(.40, .56)	(.32, .52)	(.34, .50)

Note: Numbers in parentheses represent the 95% confidence interval. MZ =

Monozygotic; DZ = Dizygotic; CTCT = Cross-twin cross-time

Table 9. *Genetic and environmental contributions to authoritarianism*

	Time 1	Time 2	T1-T2 correlation	Proportion of Influence on Stability
A	.46 (.26, .67)	.41 (.18, .61)	1.00 (.87, 1)	.58 (.34, .82)
C	.22 (.02, .39)	.19 (.00, .39)	.93 (.48, 1)	.25 (.02, .46)
E	.32 (.28, .37)	.41 (.35, .47)	.36 (.24, .47)	.18 (.11, .24)

Note: The biometric decomposition of the phenotypic covariance between T1 and T2 is represented in the last column, which shows that genetic factors account for 58% of the stability in authoritarianism. Numbers in parentheses represent the 95% confidence interval. A = additive genetic; C = shared environmental; E = nonshared environmental; T1 = Time 1; T2 = Time 2

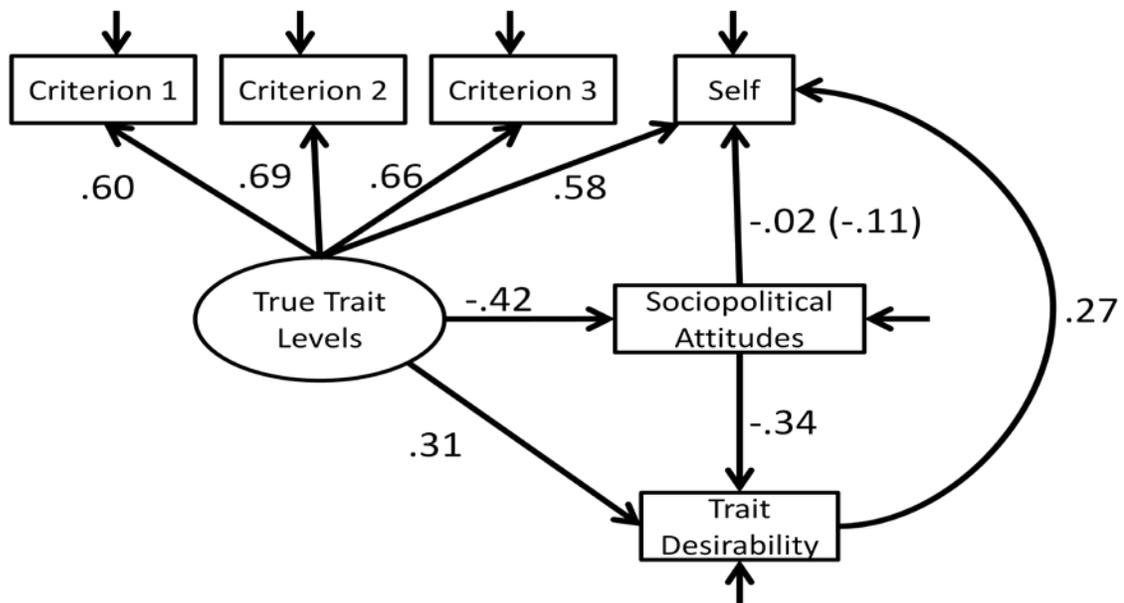


Figure 1. Mediation model to identify the influence of sociopolitical attitudes on overclaiming, by controlling for true trait variance. Parameter values shown are from the application of the model to Openness and Right-Wing Authoritarianism; for significance tests and parameters from other analyses, see Table 7. The value in parentheses represents the path coefficient in the reduced model, which omits Trait Desirability.

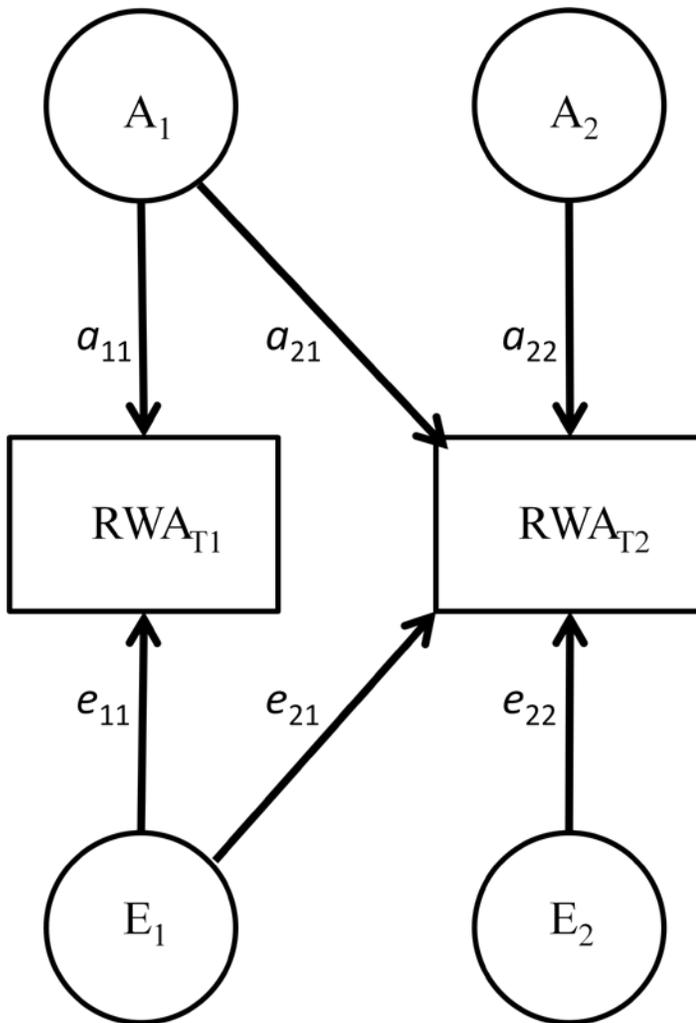


Figure 2. A path diagram of an AE Cholesky model for Right-Wing Authoritarianism (RWA) at Time 1 (T1) and Time 2 (T2). For ease of representation we have omitted shared environmental effects and represented only one member of a twin pair. In this model the variance at each time point is decomposed into additive genetic (A_1 , A_2) and nonshared environmental effects (E_1 , E_2). a_{11} and e_{11} represent additive genetic and nonshared environmental contributions to the Time 1 phenotype, respectively; a_{21} and e_{21} represent additive genetic and unique environmental contributions connecting the Time 1 and the Time 2 phenotypes, respectively; a_{22} and e_{22} represent residual additive genetic and nonshared environmental contributions to the Time 2 phenotype, respectively.

References

- Able, M. H., & Brown, L. K. (1998). Validity of the 16PF reasoning ability scale. *Psychological Reports, 83*, 904–906.
- Adorno, T., Frenkel-Brunswick, E., Levinson, D., & Sanford, R. (1950). *The Authoritarian Personality*. New York: Harper.
- Altemeyer, B. (1981). Right-wing authoritarianism. Manitoba: University of Manitoba Press.
- Altemeyer, B. (1988). *Enemies of freedom*. London: Jossey Bass.
- Altemeyer, B. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.
- Asendorpf, J. B., & Ostendorf, F. (1998). Is self-enhancement healthy? Conceptual, psychometric, and empirical analysis. *Journal of Personality and Social Psychology, 74*(4), 955–66.
- Asendorpf, J. B., & Van Aken, M. A. G. (2003). Personality-relationship transaction in adolescence: core versus surface personality characteristics. *Journal of Personality, 71*(4), 629–66.
- Berglof, E. (2011). Crisis In Transition. Retrieved from <http://www.ebrd.com/downloads/research/transition/tr11.pdf>.

- Bilsky, W., Munster, W. W., & Schwartz, S. H. (1994). Values and personality. *European Journal of Personality*, 8, 163–181.
- Blonigen, D. M., Carlson, M. D., Hicks, B. M., Krueger, R. F., & Iacono, W. G. (2008). Stability and change in personality traits from late adolescence to early adulthood: a longitudinal twin study. *Journal of Personality*, 76(2), 229–66.
- Borkenau, P., Zaltauskas, K., & Leising, D. (2009). More may be better but there may be too much: optimal trait level and self-enhancement bias. *Journal of Personality*, 77(3), 825–58.
- Caprara, G. V., Barbaranelli, C., Consiglio, C., Picconi, L., & Zimbardo, P. G. (2003). Personalities of politicians and voters: Unique and synergistic relationships. *Journal of Personality and Social Psychology*, 84(4), 849–856.
- Caprara, G. V., Barabaranelli, C., & Zimbardo, P. G. (1999). Personality profiles and political parties. *Political Psychology*, 20(1), 175–197.
- Chamorro-Premuzic, T., & Furnham, A. (2004). A possible model for understanding the personality--intelligence interface. *British Journal of Psychology*, 95, 249-64.
- Conn, S. R., & Rieke, M. L. (1994). The 16PF 5th edition technical manual. Champaign, IL: Institute for Personality and Ability Testing.
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136(6), 1092-122.

- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*(6), 1138–51.
- DeYoung, C. G. (2011). Intelligence and personality. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 711–737). New York: Cambridge University Press.
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5), 880–96.
- Duckitt, J., & Fisher, K. (2003). The impact of social threat on worldview and ideological attitudes. *Political Psychology, 24*(1), 199–222.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology, 37*(2), 90–93.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York, NY: Dryden.
- Edwards, A. L., & Walker, J. N. (1961). A short form of the MMPI: the sd scale. *Psychological Reports, 8*, 485–486.
- Egan, V. (1989). Links between personality, ability and attitudes in a low-IQ sample. *Personality and Individual Differences, 10*, 997–1001.

- Funk, C. L., Smith, K. B., Alford, J. R., Hibbing, M. V., Eaton, N. R., Krueger, R. F., ... & Hibbing, J. R. (2012). Genetic and Environmental Transmission of Political Orientations. *Political Psychology*.
- Furr, R.M., & Funder, D.C. (2007). Behavioral observation. In R. Robins, C. Fraley, & R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 273–291). New York: Guilford Press.
- Hirsh, J. B., DeYoung, C. G., Xiaowen Xu, & Peterson, J. B. (2010). Compassionate liberals and polite conservatives: associations of agreeableness with political ideology and moral values. *Personality and Social Psychology Bulletin*, 36(5), 655–664.
- Jackson, D. N., & Messick, S. (1961). Acquiescence and Desirability As Response Determinants On the MMPI. *Educational and Psychological Measurement*, 21(4), 771-790.
- Jackson, D. N., & Messick, S. (1962). Response styles and the assessment of psychopathology. In S. Messick & J. Ross (Eds.), *Measurement in personality and cognition* (pp. 129 –155). New York, NY: Wiley.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102-138). New York, NY: Guilford.

- Johnson, A. M., Vernon, P. A., & Feiler, A. R. (2008). Behavioral genetic studies of personality: An introduction and review of the results of 50+ years of research. *The SAGE handbook of personality theory and assessment, 1*, 145-173.
- Johnson, W. (2007). Genetic and environmental influences on behavior: capturing all the interplay. *Psychological Review, 114*(2), 423-40.
- Johnson, W., McGue, M., & Krueger, R. F. (2005). Personality stability in late adulthood: A behavioral genetic analysis. *Journal of Personality, 73*(2), 523–52.
- Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: its structure, functions, and elective affinities. *Annual Review of Psychology, 60*, 307–37.
- Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129*(3), 339-375.
- Jost, J. T., Nosek, B. A., & Gosling, S. D. (2008). Ideology: Its resurgence in social, personality, and political psychology. *Perspectives on Psychological Science, 3*(2), 126–136.
- Jugert, P., & Duckitt, J. (2009). A Motivational Model of Authoritarianism: Integrating Personal and Situational Determinants. *Political Psychology, 30*(5), 693–719.
- Konstabel, K., Aavik, T., & Allik, J. (2006). Social desirability and consensual validity of personality traits. *European Journal of Personality, 20*(7), 549–566.

- Krosnick, J. A., & Alwin, D. F. (1989). Aging and susceptibility to attitude change. *Journal of Personality and Social Psychology*, *57*(3), 416–25.
- Krueger, J. (1996). Personal beliefs and cultural stereotypes about racial characteristics. *Journal of Personality and Social Psychology*, *71*(3), 536-548.
- Krueger, R. F., & Johnson, W. (2002). The Minnesota Twin Registry: Current status and future directions. *Twin Research*, *5*(5), 488–92.
- Li, A., & Bagger, J. (2007). The balanced inventory of desirable responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement*, *67*(3), 525–544.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Ludeke, S., Johnson, W., & Bouchard, T. J. . J. (2013). “Obedience to traditional authority:” A heritable factor underlying authoritarianism, conservatism and religiousness. *Personality and Individual Differences*, *55*(4), 375–380. Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology*, *101*(3), 620-37.
- Mar, R. A., DeYoung, C. G., Higgins, D. M., & Peterson, J. B. (2006). Self-liking and self-competence separate self-evaluation from self-deception: Associations with personality, ability, and achievement. *Journal of Personality*, *74*, 1–32.

- Martin, N. G., Eaves, L. J., Heath, A. C., Jardine, R., Feingold, L. M., & Eysenck, H. J. (1986). Transmission of social attitudes. *Proceedings of the National Academy of Sciences of the United States of America*, 83(12), 4364–8.
- McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *The American Psychologist*, 61(3), 204–17.
- McCourt, K., Bouchard, Thomas J., J., Lykken, D. T., Tellegen, A., & Keyes, M. (1999). Authoritarianism revisited: Genetic and environmental influences examined in twins reared apart and together. *Personality and Individual Differences*, 27, 985–1014.
- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51(6), 882-888.
- McCrae, R. R., & Costa, P. T. Jr. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. Pervin, (Eds.), *Handbook of personality: Theory and research*. (3rd ed., pp. 159–182). New York: Guilford Press.
- McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology*, 29(1), 96–109.
- Meston, C. M., Heiman, J. R., Trapnell, P. D., & Paulhus, D. L. (1998). Socially Desirable Responding and Sexuality. *Journal of Sex Research*, 35(2), 148–157.

- Murray, S. L., Holmes, J. G., Bellavia, G., Griffin, D. W., & Dolderman, D. (2002). Kindred spirits? The benefits of egocentrism in close relationships. *Journal of Personality and Social Psychology*, 82(4), 563–581.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81(6), 660-679.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598-609.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). New York, NY: Academic Press.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Paulhus, D. L., Harms, P. D., Bruce, M. N., & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. *Journal of Personality and Social Psychology*, 84(4), 890-904.
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic bias in self-perceptions: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 1024–1060.

- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, *60*(2), 307-317.
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, and R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224-239). New York: Guilford Press.
- Paulhus, D. L., & Trapnell, P. D. (2008). Self-presentation: An agency-communion framework. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality psychology*. New York: Guilford.
- Pauls, C. A., & Stemmler, G. (2003). Substance and bias in social desirability responding. *Personality and Individual Differences*, *35*(2), 263-275.
- Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763.
- Pytlik Zillig, L. M., Hemenover, S. H., & Dienstbier, R. A. (2002). What do we assess when we assess a Big 5 trait? A content analysis of the affective, behavioral, and cognitive processes represented in Big 5 personality inventories. *Personality and Social Psychology Bulletin*, *28*(6), 847–858.
- Ray, J. J. (1979). A short balanced F scale. *Journal of Social Psychology*, *109*, 309-310.
- Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. *Journal of Personality and Social Psychology*, *84*(3), 582–593.

- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies. *Psychological Bulletin, 126*(1).
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313–345.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25.
- Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 375–398). New York, NY: Guilford Press.
- Rohan, M., & Zanna, M. (1996). Value transmission in families. In C. Seligman, J. M. Olson, & M. P. Zanna (Eds.), *Values: The Ontario symposium* (pp. 253–276). Hillsdale, NJ: Lawrence Erlbaum.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University.

- Saucier, G. (1994a). Mini-markers: a brief version of Goldberg's unipolar big-five markers. *Journal of Personality Assessment*, *63*(3), 506–16.
- Saucier, G. (1994b). Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology*, *66*(1), 141–154.
- Saucier, G. (1994). Mini-markers: a brief version of Goldberg's unipolar big-five markers. *Journal of Personality Assessment*, *63*(3), 506-16.
- Saucier, G., & Goldberg, L. R. (1996). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, *10* (October 1995), 61–77.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445.
- Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: a meta-analysis and theoretical review. *Personality and Social Psychology Review*, *12*(3), 248–79.
- Sibley, C. G., & Duckitt, J. (2010). The personality bases of ideology: A one-year longitudinal study. *The Journal of Social Psychology*, *150*(5), 540–59.
- Sibley, C. G., Wilson, M. S., & Duckitt, J. (2007). Effects of dangerous and competitive worldviews on right-wing authoritarianism and social dominance orientation over a five-month period. *Political Psychology*, *28*(3), 357–371.

- Simonton, D. (2006). Presidential IQ, openness, intellectual brilliance, and leadership: Estimates and correlations for 42 US chief executives. *Political Psychology*, *27*(4), 511–526.
- Sinha, R., & Krueger, J. (1998). Idiographic self-evaluation and bias. *Journal of Research in Personality*, *155*(32), 131-155.
- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: the impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology*, *101*(4), 862-82.
- Tellegen, A. (1978/1982). Brief manual for the Differential Personality Questionnaire. Unpublished manuscript. University of Minnesota, Minneapolis.
- Tellegen, A. & Waller, N.G. (2008). Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire, In G. J. Boyle, G. Matthews, and D. H. Saklofske (Eds.), *Handbook of personality theory and testing: Vol. II. Personality measurement and assessment* (pp. 254–285). London: Sage.
- Tsang, J., & Rowatt, W. C. (2007). The relationship between religious orientation, Right-Wing Authoritarianism, and implicit sexual prejudice. *The International Journal for the Psychology of Religion*, *17*(2), 99–120.

- Van Hiel, A., Onraet, E., & De Pauw, S. (2010). The relationship between social-cultural attitudes and behavioral measures of cognitive style: a meta-analytic integration of studies. *Journal of Personality, 78*(6), 1765–99.
- Van Hiel, A., Pandelaere, M., & Duriez, B. (2004). The impact of need for closure on conservative beliefs and racism: Differential mediation by authoritarian submission and authoritarian dominance. *Personality and Social Psychology Bulletin, 30*(7), 824–837.
- Verhulst, B., Hatemi, P. K., & Martin, N. G. (2010). The nature of the relationship between personality traits and political attitudes. *Personality and Individual Differences, 49*(4), 306–316.
- Werner, P. D., & Pervin, L. A. (1986). The content of personality inventory items. *Journal of Personality and Social Psychology, 51*(3), 622–628.
- Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology, 23*, 419-427.
- Wechsler, D. (1981). *WAIS-R manual*. San Antonio, TX: Psychological Corp.
- Zakrisson, I. (2005). Construction of a short version of the Right-Wing Authoritarianism (RWA) scale. *Personality and Individual Differences, 39*(5), 863–872.