**The Dissertation Committee for Simine Vazire Certifies that this is the approved version of the following dissertation:**

# THE PERSON FROM THE INSIDE AND OUTSIDE

**Committee:**

_____
Samuel D. Gosling, Supervisor

_____
Jane Richards

_____
James Pennebaker

_____
David Buss

_____
Tim Loving

# THE PERSON FROM THE INSIDE AND OUTSIDE

**by**

**Simine Vazire, B. A.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**May, 2006**

# Dedication

To paraphrase Newton:

"If I have seen far, it is because I have stood on the shoulders of a giant."

To Sam, for letting me stand on his shoulders.

# Acknowledgements

I have looked forward to writing the acknowledgements section of my dissertation almost more than I have been looking forward to getting my degree. This is my chance to try to convey the immense gratitude I feel to those people who have supported me and helped make all my dreams come true. There is absolutely no way I can do them justice, they deserve more thanks than anyone could express, and certainly more than this positive-affect-challenged girl can ever express. But I will give it a shot.

First and most of all, thank you Sam. I remember the first phone call I received from this man with a British accent inviting me to visit UT Austin's graduate program. I soon learned that he studied the personalities of Hyenas, made me drink absinthe on my first visit to Austin, and took me snorkeling in Barton Springs pool. We observed crawfish behavior and then went and devoured some at the Boiling Pot. I was immediately sold, and I have never looked back. Choosing to come to UT and work with you was the best decision in the world, and I want you to know that not only have I never regretted it for a second, but I've thanked the non-existent Lord at least a thousand times that I made the right choice. I also remember when I didn't get the NSF scholarship in my first year. Sam was more upset than I was, which actually made it a great day for me – I realized (not for the last time) how much my advisor cared about me and my career. He inspired me to prove the referees wrong and have a successful grad school career despite their criticisms. From day one Sam had complete faith in me, way more than I had in

myself. Without that I wouldn't be in the very fortunate position I am in today. Over the years, Sam's dedication to his students has never faded. Thank you for caring so deeply about me and all your students, not only for my research and career, but for my physical well-being (bringing me seaweed when I'm sick, soup when I had my wisdom teeth out), self-respect (sticking up for me when I face a challenge, like the nasty man-lady at the Groningen hotel), extracurricular life (watching and playing football, volleyball, basketball), and general happiness. You have been a wonderful advisor and friend, I hope the collaborations continue for a long, long time.

My life in Austin for the past five years has basically been a paradise from day one. In addition to having the best advisor in the world, I had a makeshift family as soon as I moved here. One of these people has been there for me from the very fist week of graduate school, without exception. Aaron, you have become as close to me as any family member, and I have come to depend on and trust you completely. You've made Austin feel like home, somewhere I never feel alone. Yet you've also pushed me outside of my comfort zone – getting me to read a poem at Quack's open mic, making me jump from the baby cliff at Lake Travis, encouraging me to get over my fears and insecurities and still being there for me when I failed to do so. You have had tremendous patience and understanding with me, and have put up with my irrationalities, critical ways, occasional self-centeredness, and all my other quirks. Thank you for always making yourself open to me, from opening your home, couch, computer, books, and food to me, to listening and giving me advice, to coming to me in times of need. It means so much to me that you trust me like I trust you. Even though I have always associated you with Austin and grad school, I know that our friendship isn't limited to a fixed area in space or time.

Other friendships have taken more time to develop. I don't remember exactly when or how you it happened, but somewhere along the way Pranj became extremely

important to me. I think it began with tennis matches, swimming in your pool, or going out to lunch and dinner every day, but in any case you became indispensable very quickly. It seems that we get closer and closer with every day that passes. Thanks for getting excited for me every time something good happened – you made up for all the lack of excitement I expressed. You even managed to bring out a smile in me every once in a while. Thanks for being so compassionate when things weren't going well. You can read me like a book – I never have to tell you how I'm feeling, you can always tell right away just be looking at me. Thank you for seeing past my superficial seriousness and having fun with me. And thanks for giving me a very rich and stimulating intellectual environment in my last two years of grad school. Even though I can never keep track of which paper or project you're working on at any given time, I have so much respect for you intellectually – you deserve all the successes I've had and more.

Other friendships have had to endure changes due to circumstances beyond our control. Matthias, I remember our first real conversation – on the plane returning from SPSP in Savannah. I was so flattered that you would consider me worthy of talking research with. Your friendship and mentoring has been so important to me. Even since you've moved away and become a big-shot professor I still think of you as my closest collaborator and one of my closest friends in the world. It has been so nice to know that our bond will survive through distance and changes. I am always so proud to tell people that I know you and that we work together. I remember walking home together from Benedict, rating each other's personalities, planning and executing the EAR feedback study – and finding out that our friendship can endure even the most stressful of work situations, sharing rooms at conferences all over the world. I can't tell you how much it meant to me that you treated me as your intellectual and social equal even when you were

clearly much wiser than me about psychology and about life. Thank you for sharing your wisdom with me without making me feel small. I look up to you more than you know.

Cindy, I owe a lot of my not-falling-apart-in-grad-school to you. You were there for me during the hardest and loneliest times. You were so unselfish with your time, your encouragement, and your ruthless destroying of Pranj and I at Boggle. In retrospect it is easy for me to see how I took you for granted and I hope you can forgive me for that. Thank you for always being honest with me, even if it meant growing apart at times. I wish we could have stayed close throughout the whole time, but I am really grateful that we continued to care for each other and that we didn't let the hiccups in our lives get the better of us in the long run. I hope I can someday return the favor and be there for you – please know that you can always count on me as someone who cares about you and wants the best for you.

Natasha, thank you for having so much faith in me as a teacher, mentor, and friend. You were the first person who really looked up to me and asked me to teach you and guide you. It is so flattering that you think I am capable of that, and I hope I haven't let you down. I also hope you know that even though I have often been in a leadership role in our relationship, you have taught me a lot and inspired me tremendously. You really are amazing. Even if it weren't for all the things you have overcome, your enthusiasm, optimism, and drive would be extraordinary. You don't let anything or anyone get you down for very long, you always manage to keep your eye on the important things in life and stick to your values. It is really touching that someone as bright as you would pick me as a friend and confidante. I can't wait to see where life takes you and be able to say I had some small part in helping you get there.

Navid, if I hadn't dedicated this dissertation to Sam, it would have been to you. Sam inspired me in to strive for excellence in the domain of personality psychology, but

you have been my source of motivation and my standard of excellence for as long as I can remember. I took German because you took German. I played sports because you did. I shape my values around things I think you value. You have been a sounding board for my ideas about research and life, and I value your opinion so much. Many of my friends have had to fill the role of person-I-look-up-to in your absence, but you are the original role model. You are pretty much perfect in my eyes.

Erik, you were with me to witness my development from naïve, inept personality-psychologist-wannabe into, well, someone who can at least pretend to know what I'm talking about by repeating things you've taught me. You know more about my own research area than I do, not to mention how much you know about all of your areas of expertise. You have a subtle way of letting other people share the spotlight when you're the one who deserves all the recognition. I have so much respect for you. Few people can be as proud of their past as I am, thanks to you. I hope you continue to share your knowledge and ideas about personality psychology and everything else with me. I also hope we stay close and continue to have those wonderful conversations that make me feel so much less alone in the world when we get to see each other every now and then. Thank you for your intellectual and emotional support through years and years of conferences, long-distance closeness, trips to Oregon, North Carolina, Vancouver Island, and along the I80-I680-I280 corridor from Davis to Palo Alto.

Danny, thank you for being my biggest fan. I don't know how you don't get tired of me and my egotistical lapping-up of your compliments without ever expressing my admiration for you in return. I assume you know that it's only because I have a hard time being nice, not because I don't have at least as much adoration for you as you express for me. You know that I think you are absolutely brilliant, and have an incredibly strong will and dedication to being an honest and solid researcher. You have survived in conditions

that I could never endure, and have never lost your faith in the value of science or human beings. I am really impressed. I also feel so flattered that you took so much interest in me in Italy and have made the incredible effort of staying close after having spent only one week together. We have managed to get to know each other so well and be so close – that is really rare for such a brief first encounter. Remember when we were in the taxi from Milan to Lake Orta? We talked about personality non-stop and it feels like we have never stopped talking since. I really feel like I found a kindred spirit in you. Your support has kept me going – you seem to have an endless source of unconditional positive regard for me. You don't know how much that means to me.

Stephen, thank you for giving this very sudden and unlikely relationship a chance. I agree that it is full of promise, and I know that if we want to we can make it work. We are truly lucky to have each other and I hope I will never take our relationship for granted. Amidst all the fortune in my life right now, all the promise my future holds, you are the one thing I am most excited about and treasure the most. I know that we will both have bright futures and I hope with all my heart that we get to share a long part of our futures together.

Maman, merci pour m'avoir soutenu tellement de fois. Tu sais toujours quoi dire pour me faire sentire mieux, meme si je ne le montre pas. Je suis tres fiere de toi, de la force que tu as dans la vie, de la compassion que tu montre pour tous le monde, et pour le faite que tu continue toujours a apprendre et te pousser a changer et te questioner toi-meme, les autres, et les conventions sociales.

Papa, j'ai beaucoup apprecie l'occasion de passer du temp avec toi et aller se ballader et parler ensemble quand tu est venu me render visite a Austin. Merci d'avoir ete supportif et de m'encourager dans ma recherche et mes etudes, d'avoir montre de

ix

l'interest a ma recherche et a mes amis, et d'avoir toujours ete chaleureux et tres gentil avec moi.

RebeccAsher, thank you for dragging me away from my work every once in a while, feeding me, entertaining me, comforting me, and letting me share in all of your joys and sorrows over the last few years. It's amazing how we went from living in the same town without ever seeing each other to feeling like you are childhood friends who will always be a part of my life.

Hani, thank you for caring so much about me. Your heart is so big it's unbelievable. I am lucky that you poured out some of your caring and friendliness to me – it is like a fountain of hope and support. You are such a smart, wonderful person with so much to be confident and optimistic about. Don't ever forget that.

Nate, thanks for sticking with me over the last nine (!) years, and most of all for your patience and friendship.

Jason, thank you for your support from the beginning. You made me feel welcome in the program, gave me someone to go to when I was lost and confused, collaborated with me when I barely knew what I was doing, and have shown that you consider me worthy of being your colleague. Thank you.

David Funder and Del Paulhus, thank you for taking me under your wings at a very critical point in my grad school career. The opportunity to work with my heroes gave me so much energy and drive to pursue personality psychology. My time in Riverside and Vancouver really shaped my identity as a researcher. I consider you two of the wisest people in the field, and when I get stuck, whether it be a big-picture issue or a methodological or statistical issue, I ask myself: WWDD?

Wim Hofstee, Bob Hogan, Bill Swann, Jamie Pennebaker, and David Buss: thank you for giving me the opportunity to be exposed to so many brilliant and varied

# THE PERSON FROM THE INSIDE AND OUTSIDE

Publication No._____

Simine Vazire, Ph.D.

The University of Texas at Austin, 2006

Supervisor:  Samuel D. Gosling

How do we discover a person's true personality? How does personality appear from the inside (i.e., to the self)? How does that differ from how personality appears from the outside (i.e., to the observer)? Given that people often see themselves differently than they are seen by others, what are the conditions under which each perspective is accurate? These questions are central to understanding who a person really is and, in turn, how much people are aware of their own and others' personalities. The goal of this dissertation is to examine these questions. I begin by providing a descriptive account of the differences between self- and other-perceptions in terms of positivity and accuracy. Specifically, in the first two studies, I compare how people see themselves to how they are seen by their friends, romantic partners, parents, and siblings (Chapter 2). Then, in the next two studies, I test the accuracy of self- and other-predictions of behavior by comparing them to actual naturalistic behavior recorded from people's everyday lives (Chapter 3). Finally, in the fourth study, I examine the accuracy of self, friend, and stranger ratings of personality by comparing personality judgments to laboratory-based behavioral tests of personality (Chapter 4). The results show that self-perceptions are

more negative than others' perceptions of them, people are more aware of their own negative traits than their positive traits, and they fail to notice a substantial number of their own characteristics. Observers agree substantially about what a person is like, and their knowledge of a target's observable personality is quite good. By comparing perceptions of the person from the inside and outside with objective behavioral criteria, we can come to understand the strengths and limitations of each perspective. In fact, the two perspectives often complement each other – one filling in the gaps left by the other. Furthermore, even when both perspectives are accurate, they are often accurate in different ways. Thus, although neither perspective alone can explain the whole puzzle of who a person really is, they both provide different pieces of the puzzle and together deepen our understanding of the person.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1: Introduction

In the Taiwanese film "Yi-Yi," Yang-Yang, a five-year old boy, takes Polaroids of the back of his friends' heads to show them the side of themselves that they can't see. As implied by this story, just as there are physical aspects of a person that are unavailable to the self, there are also aspects of our personalities that we do not see in ourselves. This raises troubling questions about the nature of personality and identity. Who knows a person best? What do others know about us that we don't know about ourselves? How do we uncover a person's true personality? The purpose of this dissertation is to explore such questions by comparing how people see themselves and how they are seen by others with other indicators of what they are like (e.g., how they behave).

Many people, including many personality researchers, assume that self-perceptions are almost perfectly accurate. If you want to know what a person is like, just ask them. Indeed, a reviewer once wrote "the best criterion for a target's personality is his or her self-ratings […]. Otherwise, the whole enterprise of personality assessment seriously needs to re-think itself." (anonymous reviewer, personal communication, May 5, 2003). This view, based on intuition, is empirically testable. Perhaps the field does indeed need to rethink itself. How accurate are self-perceptions? Could others' perceptions ever be more accurate? Rather than accept self-views as the "truth," I will examine the accuracy of self- and other-perceptions and the conditions under which accuracy and inaccuracy arise.

This dissertation focuses on the fundamental question of how to define who a person really is. To do this, I examine the person from the inside (the self's perspective) and outside (others' perspective), and compare the two to objective measures of what the person is like. The literature on the self and on interpersonal perception suggests that self-

perception is a fundamentally different process than other-perception, with its own unique obstacles. Thus, the goal of my dissertation is to develop our understanding of the differences between self- and other-perception. Is the self just another observer, or does the self have a unique perspective (and unique biases)? How does self-perception differ from other-perception, and under what conditions is each more likely to be accurate? In the present chapter I provide a review of the relevant literature pertaining to these issues.

## LITERATURE REVIEW

### Who should own the definition of personality?

In some sense, the question of how to define who a person really is is a philosophical one. Each researcher can choose to define personality in whatever way they want. An argument can even be made that there is no underlying reality. According to this argument, personality is in the eye of the beholder. However, decades of personality research show not only that we frequently and easily ascribe personality traits to ourselves and others, but that these traits reflect real consistencies in patterns of thinking, feeling, and behaving (Bem & Allen, 1974). Of course people are not perfectly consistent. Behavior can change as a function of external situational characteristics, a person's mood, others' effect on the person, and many other variables. Nevertheless, people have core dispositions that exert varying levels of influence in every situation. People's global personality affects their personality in specific contexts, their behavior, and fluctuations in behavior (Fleeson, 2001; Heller, Watson, & Ilies, 2044; Slatcher & Vazire, 2006). Based on these findings, I begin with the assumption that personality exists, that there is an underlying truth to be discovered about a person's character. Even with this assumption, however, it is not clear who is in the best position to describe a

person's true personality. There are theoretical arguments in favor of both the self and others' perspective, which I review below.

Ultimately one's definition of personality will depend on one's reasons for assessing personality. In the research context, the primary purpose of personality assessment is usually to predict future outcomes or aggregated (or repeated) behaviors. For example, we may want to predict employee productivity from conscientiousness. Thus, I will use behavioral outcome measures as the criterion whenever possible. As a consequence my research focuses heavily on the behavioral aspects of personality, with less attention to the cognitive and affective components of personality. This is consistent with my view that although personality encompasses many non-behavioral characteristics, all aspects of one's personality eventually manifest themselves in behavior. In other words, according to this view one cannot possess a personality trait that does not manifest itself in any observable behavior.

Although this view is controversial, it is not without precedent. Buss and Craik's (1985) act frequency approach to personality is based on the same view of personality. This approach does not preclude me from examining traits that are cognitive (e.g., intelligence) or affective (e.g., anxiety) in nature, but it implies that even these traits can be assessed through behavioral observation. I define behavior broadly to include test responses (e.g., IQ test scores), laboratory behavior (e.g., participation in a group exercise), long-term outcomes (e.g., staying in a relationship or breaking up), and even physiological markers (e.g., testosterone levels) as well as spontaneous behavior in everyday life (e.g., laughing).

Alternative approaches include Costa and McCrae's Five Factor Model (McCrae & Costa, 1999), which proposes that traits stem from basic tendencies (which have a biological basis) that, in conjunction with environmental conditions and dynamic

3

processes between biology and the environment, produce characteristic adaptations, which are essentially traits (patterns of thought, feeling, and behavior). Another model, the Cognitive-Affective Personality System (CAPS; Mischel & Shoda, 1999) proposes that behaviors are generated from a cluster of affective and cognitive units that interact with each other and react to situational features. While I do not necessarily disagree with these models, I find them difficult to test and impractical to implement in empirical research. For these reasons, and because I am interested in applying my research to behavioral prediction, I have chosen to base my work on the more behaviorally-oriented act frequency approach to personality.

Although my decision to focus primarily on behavioral aspects of personality is theoretically motivated, it also provides a practical solution to the most daunting problem in personality research – the criterion problem. The question of accuracy in personality judgment has been avoided by many researchers because of the problem of identifying a criterion for accuracy – how do we gauge whether a perception of someone's personality is correct? If I say I am extraverted, by what standard do we evaluate the accuracy of my self-judgment? One strong candidate for a criterion is behavior – I am correct if my judgment corresponds with my behavior.

Behavioral measures are far from perfect indicators of personality, but they have several important advantages. I review the strengths and limitations of behavioral measures of personality below. For behavior to fulfill its role as an objective criterion, it must be assessed independently of self- and other-reports. That is, we must measure behavior directly and not ask the person or his or her friends to report on their behavior (Block, 1989). As demonstrated in the studies I present, my desire for objective, ecologically valid measures of behavior has led me to search in many unusual places and to make use of a broad range of methods.

Another reason for the dearth of research on accuracy is the difficulty in obtaining a valid statistical test of accuracy. Cronbach (1955) raised several important statistical issues that have stood in the way of obtaining valid indices of accuracy. Current statistical tools for assessing overcome many, but not all, of these obstacles. For example, comparing the mean levels of a judgment to a criterion can tell us about the amount of elevation (i.e., positivity or negativity bias), but does not tell us about whether the judge knows about the target's relative standing on different traits. In contrast, correlational measures of accuracy can tell us rank-order accuracy across traits but not about elevation. Thus, to address some of these concerns, it is necessary to compare judgments and criteria using several different indices (e.g., mean levels and correlations). I used this approach whenever appropriate throughout my dissertation. In the following sections, I examine the conceptual and empirical arguments related to the accuracy and inaccuracy of self- and other-perceptions of personality.

**The Person from the Inside: Self-Knowledge**

*Conceptual Issues*

There are several conceptual reasons to expect that people have a great deal of insight into their own personality. The first reason to expect that the self has an advantage in assessing personality and predicting behavior is that people are exposed to more of their own behavior than anyone else. This leads not only to a greater quantity of information about their behavior, but also a more diverse range of information – they have witnessed their own behavior in more situations than anyone else. According to Bem's self-perception theory (Bem, 1967), this would present people with an advantage over others in making inferences about their own personality.

There are several reasons to question this advantage, however. The first is simply a matter of memory capacity – can we really retain all of that information? Our close friends and family are also exposed to many instances of our behavior, and it is possible that limits on memory capacity make differences in amount of information between self and close others irrelevant. The second reason to doubt that the self has privileged access to its own behavior is that we are not simply neutral observers of our behavior – we are busy trying to decide how to act and react, and so cannot devote all of our cognitive resources to assessing our own personality. Research has supported this claim by showing that people pay more attention to situational influences when they are in the 'actor' position than in the 'observer' position (Jones & Nisbett, 1971; Nisbett & Ross, 1980). This lack of attention to dispositional influences may lead us to underestimate the cross-situational consistency of our behavior, and hence miscalculate our own personality traits. Nevertheless, there is probably some truth to the claim that our broad range of exposure to our own behavior offers some advantage in personality assessment. For example, college students probably have a unique perspective on their personality because they know how they act with both their college friends and their family, whereas their friends and family typically know little about how they act with the other (Funder, Kolar, & Blackman, 1995).

The second reason to expect that the self has an advantage in assessing personality and predicting behavior is that only the self has direct access to internal states. It is reasonable to assume that knowing one's thoughts, feelings, and motivations will lead to greater accuracy in description and prediction of behavior. For example, if Tim went to a party and behaved socially, but he knows he only did it to impress a girl and felt very uncomfortable doing it, he will probably claim, accurately, that he is not extraverted and that he is not likely to act extraverted at future parties. Tim's friends,

however, are more likely to think that Tim was being himself and that his behavior was diagnostic of his personality. One flaw in this argument is that Tim's good friends will probably see Tim often enough to detect his introverted nature. If they do not, this implies that Tim consistently acts extraverted. In this case, is it more accurate to say that Tim is introverted, because that is how he sees himself, or that he is extraverted, because that is how he behaves? Using the criterion established earlier, behavioral prediction, Tim's friends' assessment would be more accurate and useful in predicting behavior. That Tim sees himself as introverted is relevant to his self-concept, but not to his standing on the trait of extraversion. The self's privileged access to internal states does confer it an advantage in some domains of personality assessment. Specifically, knowledge about internal states is particularly likely to be useful in assessing aspects of behavior that have low visibility and a large subjective component, such as affective traits, goals, fantasies, and intentions. This possibility has been tested in previous research, described below.

There are also important reasons to doubt people's ability to provide accurate self-reports. First, self-report questionnaires are often not as straightforward as they seem (Schwarz, 1999). Research has shown that small differences in how the questions are worded can have a large impact on people's responses. Second, people may simply choose not to tell you. Impression management, or the desire to present oneself in a specific, inaccurate way, can lead to many response biases, such as socially desirable responding, false humility, and so on. Impression management is especially likely to occur in individuals or situations which are very sensitive to situational demands (e.g., job interviews). Other types of response bias may also affect the quality of self-reports, such as acquiescence or the extremity response bias.

Another important potential bias in self-reports is people's persistent, inaccurate views of themselves. In many cases, these errors stem from self-deception – people

ignore, distort, or reinterpret information that does not fit with their existing or preferred self-view (e.g., Gosling, John, Craik, & Robins, 1998; Robins, Noftle, Trzesniewski, & Roberts, 2005). Whereas impression management reflects a conscious bias on the part of the respondent, self-deception is unconscious. Many argue that self-deception should not be controlled for in self-reports because they reflect a real aspect of personality (McCrae & Costa, 1983; Paulhus & John, 1998). However, if the researcher's concern is not to measure self-deception but to obtain accurate ratings of personality traits, self-deception poses a serious threat to validity.

It is often assumed that an honest self-disclosure, untainted by impression management or self-deception, would yield an accurate self-description. The assumption is that the truth is fully available to the self. However, there is good reason to believe otherwise (Paulhus & Vazire, in press). For one thing, some information is unavailable to the self-reporter. Dunning and colleagues (2005) distinguish between information unavailable to the self-assessor and information that tends to be ignored by the self-assessor. In addition, people do not have an infinite ability to recall all information relevant to a posed question. Conversely, they may be over-whelmed with an abundance of information. Its integration and simplification may be too challenging a task. Self-reporters may have to resort to a "press release" version of their personality just to get on with the task.

Even worse than lack of access is the claim that introspection may actually diminish accuracy. According to Wilson and Dunn (2004) any extended attempt to clarify one's self-descriptions can undermine their validity. I know of no such research on the validity of personality self-reports. However, the test administrator's preference for superficial response is evident in common instructions such as: "give your first answer"; "work quickly"; "do not spend too much time any one item".

Nevertheless, self-perceptions represent an important aspect of personality. To the extent that the self does have unique insight and information, self-reports of personality are uniquely accurate. Furthermore, self-perceptions reflect a person's phenomenological experience of their own personality, which is a vital aspect of who a person is (McAdams, 1996).

### *Individual Differences*

Research has shown that some people are more prone to errors and biases in self-reports of personality and behavior than others. John and Robins (1994) conducted a study in which participants ranked their own and others' performance on a group managerial task. In addition, they were ranked by an assessment team of psychologists, and their self- and other ratings were compared to this criterion. Although the results showed an overall self-enhancement effect, the data also revealed strong individual differences ranging from self-enhancement to self-diminishment, and these differences were predicted by narcissism measures.

In another study, Robins and John (1997) replicated this result using the same paradigm, finding that narcissists enhanced their own performance more than nonnarcissists. In addition, they also found that changing the participants' visual perspective (having them watch a video of themselves in the group interaction) increased self-enhancement in narcissists but had the opposite effect in nonnarcissists. Finally, a study by Gosling, John, Craik, & Robins (1998) using a similar paradigm found that there were vast individual differences in self-enhancement when reporting act frequencies. Specifically, Gosling et al. found that 57% of participants showed self-enhancement, 24% showed self-diminishment, and 19% were accurate, and narcissism correlated .27 with self-enhancement. Together, these findings show that narcissists' self-reports are particularly biased and immune to situational influence, whereas nonnarcissists' self-

reports tend to be more realistic and can be improved by changing the person's visual perspective. Below I present empirical research evaluating the accuracy of self-reports.

*Empirical Research*

In the most direct comparison of the relative accuracy of self- and peer reports against a behavioral criterion, Kolar, Funder, & Colvin (1996) used self- and informant ratings to predict participants' behavior in a short videotaped interaction with an opposite-sex stranger. Although their results overwhelmingly support the accuracy of informant ratings as opposed to self-ratings, the results nevertheless provide some interesting patterns of accuracy in self-ratings. Specifically, self-ratings outperformed informant ratings in predicting the following behaviors (9 out of 41 California Q-Sort/Behavioral Q-Sort items): "appears to be relaxed and comfortable," "is reserved and unexpressive," "expresses skepticism or cynicism," "exhibits high degree of intelligence," "expresses sympathy towards partner," "expressive in face, voice, or gesture," "discusses philosophical issues with interest," "behaves in masculine/feminine style," and "speaks quickly." Many of these traits are related to affect – the expression of emotion. These data suggest that although informant reports of personality have greater accuracy when predicting behavior, the self's unique access to internal states, including emotional experience, leads to greater accuracy in predicting emotional expression.

Another study by Spain, Eaton, and Funder (2000) followed up on this issue by explicitly comparing the ability of self- and informant reports to predict both behavior and emotional experience. In this study, self- and informant reports were used to predict two criteria: the participants' emotional experience in everyday life and the participants' behavior in the laboratory. Emotional experience was measured using the experience sampling method – participants wore a beeper that signaled them four times a day for eight days and rated their positive affect and negative affect when signaled. The

behavioral measures were the same as in the study described above. The results from the study by Spain, Eaton, and Funder support the earlier finding that self-reports are more accurate than informant reports in predicting affective traits. Specifically, self-reports of extraversion and neuroticism predicted positive affect and negative affect respectively, whereas informant reports only predicted positive affect, and the effect was weaker for informant reports than self-reports. Surprisingly, self-reports of extraversion also predicted extraversion-related behavior in the laboratory better than informant reports. This effect is inconsistent with Kolar et al.'s (1996) earlier findings, and may be due to the fact that Spain, Eaton, and Funder's study restricted personality assessment to the traits of extraversion and neuroticism. In addition to this limitation, this study also fails to provide indisputable evidence for the accuracy of self-reports with respect to affective traits because of the methodology employed. The use of experience sampling as measure of positive and negative affect may have artificially increased the predictive ability of self-reports of personality, because experience sampling is itself a self-report method. Nevertheless, this study does provide strong evidence for the ability of self-ratings to predict emotional experience, if not emotional expression.

In another study with an objective behavioral criterion, Borkenau and Liebler (1993) examined the accuracy of self-, acquaintance, and stranger ratings of personality and intelligence. The personality data will not be described here because it does not shed any light on the relative accuracy of self- and informant reports, but the intelligence ratings were compared against measured intelligence and revealed an interesting pattern. Correlations with measured intelligence were .32 for self-ratings, .29 for acquaintance ratings, and .43 for stranger ratings (based on a short video clip of the target reading a weather report). These results suggest that although people do have some insight into

their objective level of intelligence, they do not have more information than their friends or even someone who has merely watched them read a weather report.

Finally, Gosling et al. (1998) examined the accuracy of self-reported act frequencies when compared to observer codings of act frequencies. Their research employed a videotaped group managerial task and asked participants to report immediately after the interaction the frequencies with which they performed various acts during the interaction. Their results indicate that the accuracy of self-reports varies greatly across acts and have an average correlation with observer codings of .26 for all acts and .40 for acts coded reliably by observers. The pattern of results showed that self-reports were more biased for more desirable acts, supporting the theory that there are motivational forces that contribute to the inaccuracy of self-reports.

The results from these studies cast some doubt on the ability of self-reports to predict behavior. The most promising trend in these results, however, is that self-reports may be a useful method for assessing affective traits, particularly those related to emotional experience and possibly extending to emotional expression.

I next turn to studies of self-other agreement to evaluate the accuracy of self-reports. Although moderators of self-other agreement do not tell us directly about the relative accuracy of self- and informant reports, careful interpretation of the data can provide us with some possible explanations for the moderating variables that may give us some insight into the conditions of self-report accuracy.

Consistent with the evidence presented above on behavioral measures of the accuracy of self-reports, Funder (1980) found patterns of self-other agreement that suggest that self-reports are more accurate when the ratings pertain to internal states. Specifically, the results showed that people gave themselves higher scores on internal attributes (e.g., "is introspective" and "engages in fantasy") than did their friends, but

gave themselves lower scores on external attributes (e.g., "is an interesting person" and "is personally charming") than did their friends. The most likely explanation for this finding is that the self's privileged access to internal states comes at the expense of attention to external behaviors. This closely parallels the finding that self-ratings predict affect-related behaviors and emotional experience more accurately than informant reports, whereas they are inaccurate at predicting most other behaviors.

In another illuminating study on self-other agreement, Pelham & Swann (1994) found that there was more convergence between self- and other ratings on those traits about which targets reported high levels of certainty. The most likely explanation for this finding is that people were less accurate in their self-ratings on traits about which they were less certain. In other words, people were accurate when they reported being certain, and inaccurate when they reported being uncertain. Similarly, Bem and Allen (1974) found that self-rated consistency on a trait leads to greater self-other agreement, and Cheek (1982) found that self-reported importance of personality traits predicted self-other agreement on these traits. This promising finding suggests that simply getting certainty ratings for each item on self-reports could tell us a lot about how much confidence we should place in the accuracy of each of those ratings. Sometimes the answer is as simple as asking people.

As Hayes and Dunning (1997) demonstrated, trait ambiguity is another variable associated with impaired self-other agreement. In their study, ambiguous traits (e.g., sophisticated) led to lower self-other agreement than unambiguous traits (e.g., punctual). A complex regression analysis showed that this effect was especially pronounced in the desirable traits, suggesting that ambiguity contributed to an enhancement effect in self-ratings.

In an unusual study, Davidson (1993) dealt a blow to the accuracy of self-reports by showing that self-other disagreement is related to behavioral measures of thought suppression and repression. In other words, those people who disagreed most with their friends about their own personality were also the ones who were most successful at suppressing unwanted thoughts (e.g., the white bear paradigm). This evidence is admittedly indirect at best, but raises some interesting questions about the nature of self-other disagreement and suggests a possible mechanism for self-enhancement.

Finally, John and Robins (1993) addressed several determinants of self-other agreement, some of which have implications for the accuracy of self-reports. As suggested by the research described above, there is some evidence that self-reports are particularly inaccurate for highly evaluative traits. John and Robins's research supports this by showing that self-other agreement is highest for neutral traits and lowest for highly desirable and highly undesirable traits. Other research (e.g., Gosling, John, Craik, & Robins, 1998) has reported inconsistent evidence, demonstrating a desirability effect such that self-other agreement is *highest* for desirable traits and lowest for undesirable traits. This discrepancy, however, may be due to the fact that this research did not include many undesirable or even neutral traits, whereas John and Robins (1993) included traits with a wide range of desirability. The Hayes and Dunning (1997) study described above also sheds some light on this issue. Their results show that among friends, there is a desirability effect (i.e., higher self-other agreement for desirable traits than neutral or undesirable traits) but among non-friend roommates, there is an evaluativeness effects (i.e., higher self-other agreement for neutral traits than desirable traits). All of these findings should be interpreted with John & Robins's (1994) and Robins and John's (1997) research in mind, indicating that individual differences in narcissism may be moderating self-enhancement effects.

Another interesting finding of the John and Robins (1993) study is that self-other agreement was actually highest for the trait 'Emotional,' contrary to the privileged access hypothesis described above and supported by Spain et al. (2000). It is therefore unclear whether people have any advantage over their close friends even when rating their own internal traits.

**The Person from the Outside: Informant Reports**

*Conceptual Issues*

Although the idea of using informant reports for assessing personality is less intuitive than using self-reports, there are actually more conceptual reasons for expecting informant reports to be accurate. The first and best reason was stated most clearly by Hofstee (1994): "There is only one me, whereas there are many others who know enough about me to provide a more reliable average judgment. Thus, other things being equal, I am outnumbered and outperformed by the average other." The principle of aggregation is clear-cut – an increased number of judges leads to a more reliable assessment, and since there is only one self, only informant reports can reap this benefit.

Another benefit informants have in assessing personality is the advantage of perspective. As the actor-observer literature has shown (e.g., Jones & Nisbett, 1971), observers tend to pay more attention to dispositional influences on behavior than actors. Presumably, this will lead observers' judgments of personality to be more accurate than actors'. Observers' physical perspective also gives them greater ability to directly observe the target's behavior than the self has. According to looking-glass self theory (Cooley, 1902; Vazire, in press-b), self-perceptions are based on reflected appraisals of how we think others see us. If this is true, informants' perceptions are a better measure of personality because they are based on direct observation rather than reflected appraisals.

15

Furthermore, informant reports are a measure of a person's reputation (Hogan, 1998). Reputation is an important indicator of personality because reputations are based on past behavior, and as we know, the best predictor of future behavior is past behavior. In addition, reputations have an important causal force in people's lives – what others think of us can influence how they treat us, what opportunities they give us, what they tell others about us, etc..

Finally, the most obvious benefit of informant reports is that they are much less susceptible to influence by self-deception and self-presentational motives. Of course it is likely that people will enhance their close friends, but this effect is generally much weaker than self-enhancement and self-diminishment effects (John & Robins, 1993).

All of these arguments speak to the potential of informant reports to serve as a robust and valid method of assessing personality. Below I address the question of who these informants should be, and then review the evidence for the accuracy and inaccuracy of informant reports. I will revisit much of the research described above in the discussion of the accuracy of self-reports, but this time will focus on the findings that reflect on the accuracy of informant reports.

### Individual Differences

As discussed above, the characteristics of the target being judged are an important factor in determining the validity of personality assessment methods. Namely, narcissists do not provide valid self-reports of their personality. Similarly, the characteristics of the informant providing personality ratings are an important factor to consider before measuring the accuracy of informant reports.

Who makes a good judge of personality? There is very little empirical data giving a straight answer to this question, but research on self-other agreement suggests some possibilities. One variable that has been found to influence self-other agreement is the

level of acquaintance between the informant and the target (Watson, Hubbard, & Wiese, 2000). For example, Paulhus and Bruce (1992) found that acquaintance over a seven-week period increased self-other agreement from .21 (at week 1) to .30 (at week 7). Other research (e.g., Borkenau & Lieber, 1993) has shown that the effect of acquaintanceship on self-other agreement varies across traits, with Extraversion being the least affected by acquaintanceship. Still other research has found that cohabitation is a stronger influence on self-other agreement than acquaintanceship (Bernieri, Zuckerman, Koestner, & Rosenthal, 1994). Finally, Watson, Hubbard, and Wiese (2000) found a stronger acquaintanceship effect when measuring the accuracy of informant ratings of emotional traits than other traits, suggesting that emotional traits may take longer to manifest themselves to observers than traits such as extraversion.

Another factor to take into consideration is the relationship between the informant and the target. As Swann (1984) proposed, personality judgments are embedded in an interpersonal context that is shaped by the nature of the interpersonal relation and the purpose of the personality judgment. Specifically, Swann's view posits that people form personality impressions of others that are highly accurate in the circumscribed context of their interactions with the target. Rather than present this as a limitation on accuracy, Swann shows that researchers should reconceptualize accuracy to account for the pragmatic concerns of informants. Thus, in order to bolster the validity of informant reports, it follows that researchers should obtain informant reports from multiple informants who know the target in multiple contexts. Furthermore, those who know the target across a wider range of contexts will probably develop a more nuanced assessment of the target's personality.

Funder, Kolar, and Blackman (1995) have addressed this by looking at the agreement among informants from three different contexts; hometown friends, college

17

friends, and parents. Consistent with Swann's model, their results showed that informants in each context had formed slightly different impressions of the targets – agreement among judges was higher within contexts than across contexts (.33 for hometown friends, .46 for college friends, and .42 for parents). It is important to note, however, that there was significant agreement among judges across contexts as well, which suggests that the aggregate of informant reports across multiple contexts would provide the most reliable set of informant reports. Together, these findings demonstrate that informant reports should consist of ratings from multiple judges who are well acquainted with the target across a variety of contexts.

### *Empirical Research*

In the Kolar, Funder, and Blackman (1996) study in which they used self- and informant reports to predict behavior in a short, videotaped interaction, the overall results demonstrate a significant advantage for informant reports. Specifically, the average correlation between the combined informant reports and behavioral codings was .18, significantly higher than the .12 average correlation between self-reports and behavioral codings. This effect was stronger when the behavioral codings were aggregated to form four behavioral factors. Across these four factors, the average correlation between informant reports and behavioral codings was .44, whereas the average correlation between self-reports and behavioral codings was .28.

Consistent with these findings, John and Robins (1994) found that in a group managerial task, participants agreed more with the psychologists' ratings of performance when evaluating their peers (mean $r$ = .84) than when evaluating themselves (mean $r$ = .58).

In perhaps the most impressive study of the accuracy of informant reports, Levesque & Kenny (1993) used a round robin design to examine the relative accuracy of

self- and informant ratings in predicting subsequent behavior in dyadic interactions. The participants were briefly introduced to one another and then rated each participant's personality and made behavioral predictions. Their results demonstrate that even after such brief exposure to targets, unacquainted informants were better at predicting targets' behavior than were the targets themselves (average r between informant reports and coded behavior = .65, average r between self-reports and coded behavior = .39).

In contrast to these studies, Spain et al. (2000) found little evidence for the accuracy of informant reports in predicting laboratory behavior or emotional experience. Again, this study was limited to the traits of extraversion and neuroticism, and used experience sampling as the method for measuring emotional experience. These limitations, in addition to the body of evidence presented in contrast to these results, suggest that these findings do not reflect the general accuracy of informant reports in predicting behavior. On the contrary, the evidence presented here support Hofstee's (1994) view that aggregated informant ratings are more accurate than self-ratings in assessing personality. More research is needed on differences in the validity of informant ratings across traits to determine whether Spain et al.'s (2000) findings are due to the nature of the traits measured, and whether informant ratings are less accurate for extraversion and neuroticism than for other traits.

There have also been studies of the accuracy of self- and informant-reports in the clinical literature. Specifically, research by Oltmanns and his colleagues suggests that peer-ratings of personality disorders are extremely useful for several reasons. First, peer-nominations can be aggregated and so are more reliable than self-reports (Oltmanns, Turkheimer, & Strauss, 1998). Second, certain types of personality disorders are characterized by a lack off awareness of the problem (e.g., narcissistic personality disorder), and thus, extreme cases cannot be captured with self-reports (Oltmanns,

Gleason, Klonsky, & Turkheimer, 2005). Third, peer-nominations predict outcomes such as early discharge from the military and provide incremental validity above and beyond self-reports (Fiedler, Oltmanns, & Turkheimer, 2004).

In addition to these direct tests, John and Robins's (1993) review of determinants of self-peer agreement revealed several patterns supporting the greater accuracy of informant reports over self-reports. As described above, their results showed an effect for evaluativeness such that self-peer agreement was higher for neutral than for desirable or undesirable traits. Interestingly, this pattern was less pronounced for peer-peer agreement (peer-peer agreement was only slightly higher for neutral traits), indicating that informant reports are less susceptible to enhancement and diminishment biases than self-reports.

This same study also found an effect for trait observability, such that self-peer and peer-peer agreement was higher for more observable traits, suggesting that informant reports are more valid for more observable traits. This finding has been replicated by several other researchers (e.g., Hayes & Dunning, 1997).

Consistent with previous findings on the moderating effect of observability, Watson, Hubbard, and Wiese (2000) found that self-other agreement was higher on the Big Five traits than on affective traits. The results showed that this was probably due to an effect of assumed similarity for affective traits – informants tend to rate targets as similar to themselves on affective traits. This is probably a direct consequence on the lack of information available to informants for affective traits, which are low in observability.

In addition, Hayes and Dunning (1997), demonstrated an important benefit of aggregating informant reports. In their study of trait ambiguity as a moderator of interjudge agreement described above, their results showed that aggregating informant reports helps to attenuate the effect of ambiguity on the accuracy of personality assessment. This is because aggregation reduces error due to the idiosyncratic

20

interpretation of ambiguous traits, so that when multiple informants are used, the validity of personality assessment is preserved, even for ambiguous traits.

Finally, there is evidence from several studies that informants are aware of these moderators of accuracy. For example, Funder and Dobroth (1987) found that agreement among informants across traits was positively correlated with subjective judgments of trait visibility by these same informants. In other words, informants are sensitive to the difference in visibility between traits, and their level of accuracy in personality assessment corresponds to their perceptions of trait visibility. In another study, Fuhrman and Funder (1995) found that the extremity, frequency, and speed with which informants endorsed specific traits when rating the target was correlated with the target's self-schema (as measured by the extremity, frequency, and speed with which the target endorsed the traits). Together these studies suggest that people are aware of what they know and don't know about a target – they can discriminate between traits with high and low visibility, and they respond more quickly and extremely to traits that are central to the target's self-schema. In order to confirm this trend, however, it would be useful to conduct a study in which researchers measure the informants' level of certainty for each trait rating, and examine whether the certainty ratings correspond to the accuracy of the trait ratings (as measured against a behavioral criterion). Evidence from self-reports presented above shows that certainty ratings are a useful measure of accuracy for self-reports, but the same has not been tested directly for informant reports.

**Criterion Measure: Behavior**

The evidence reviewed thus far suggests that both self and informants are promising sources of information about a person. Many of the studies described have grappled with the problem of how to assess the accuracy of these perspectives. As McCrae is quoted as saying in Hofstee's (1994) paper, ultimately "it is reality that owns

the truth about personality, not any single operationalization." The challenge, and the goal of this dissertation, is to compare self-views and informants' views to reality and find out who knows what about a person. In this section I review several methods for assessing this elusive reality, namely, behavioral measures or personality.

### *Conceptual Issues*

The appeal of behavioral measures runs deep in the social sciences. In both animal and human research (e.g., in anthropology and psychology) there is a widespread assumption that behavioral measures are more objective, and thus a better reflection of reality, than global trait ratings (e.g., self or informant reports). Among personality researchers, there are several possible reasons for this assumption. First, to researchers who accept the behavioral approach to personality outlined above, behavior is the essence of personality. That is, extraversion is defined as acting extraverted, regardless of how that behavior is experienced by the actor. Second, behavior is, at least in theory, an objective criterion. It can be used to assess the accuracy of both self- and other-judgments of personality. For example, if I say I am extraverted but my friends say I am introverted, we can use a behavioral criterion to determine which judgment is more accurate. Many researchers assume that behaviors are less subjective and more easily interpreted than trait ratings.

Clearly measuring personality-relevant behavior directly is an appealing approach. However, as anyone who has actually coded behavior knows very well, behavior codings can quickly become very complex (Vazire, Gosling, Dickey, & Shapiro, in press). Subjectivity is introduced in the decisions that need to be made both during the development of the coding system and its implementation. The validity of behavior codings can easily be compromised if researchers do not interpret behaviors correctly, if they attribute the behaviors to incorrect traits, or, more likely, if they simply

neglect to account for the fact that a single behavior can be indicative of several different traits

First, researchers must decide which behaviors to include when developing the coding system. This decision involves determining which behaviors are indicators of which traits. For example, researchers dictate the meaning of a category, such as "abnormal behavior," by deciding which behaviors should be included in the category and which ones should not. Selecting behaviors that are psychologically meaningful and categorizing them is a process that is subjective and susceptible to error.

Next, researchers must decide at what level to code the behaviors. The most micro level of analysis might be recording specific muscle movements. This has the advantage of being precise and objective, but makes subsequent steps of collecting and interpreting data much more difficult. As Martin and Bateson (1993) wrote: "the cost of gaining detail can be that higher-level patterns, which may be the most important or relevant features, are lost from view." (p. 9). In contrast, researchers could choose to measure behavior at the most global level, such as coding psychologically meaningful acts (e.g., arguing). As I discuss below, this introduces problems in the implementation stage.

In addition to the subjective decisions that are made in designing the coding system, subjectivity is also introduced when the behavior-coding system is implemented. For example, observers may apply the behavioral definitions strictly or loosely. As Block (1989) has noted, if the behavioral definitions are applied strictly--which presumably they must if the feared subjectivity is to be avoided--then relevant behaviors could easily be missed and irrelevant behaviors captured.

Another way in which the implementation of behavior-coding systems is subjective is that many behaviors are ambiguous. This is particularly true if the behaviors are defined at the psychologically meaningful macro level. Even carefully defined

23

behaviors are not immune from interpretational issues. For example, walking seems like a very simple, easy-to-define behavior, but what counts as walking? One step? Two? Another obstacle to using behavioral measures of personality is the fact that behavior is very difficult to measure reliably. Whereas self- and informant-ratings can easily be aggregated using multi-item scales or multiple informants, it is difficult to obtain reliable, multi-item measures of behavior. One reason is the sheer impracticality of obtaining behavioral measures in humans. Measuring behavior in people's natural environments is almost impossible (although there are exceptions which I discuss below), and humans are only willing to stay in a lab for a limited amount of time. Another reason for the low reliability of behavioral measures is that they are incredibly sensitive to situational fluctuations, but do not take these contextual factors into account. For example, someone shouting at their friend to warn them about a bee and someone shouting for fun at a party are both performing the same behavior, but the behavior reflects different personality characteristics in the two contexts. Behavioral measures do not take the contexts into account and so are less stable than the actual underlying personality traits producing the behaviors.

Despite these obstacles to obtaining objective, reliable behavioral measures, behavior is still arguably the best source of information we have for gauging the accuracy of self and informant reports. Winston Churchill once said "No one pretends that democracy is perfect or all-wise. Indeed […] democracy is the worst form of government, except for all those other forms that have been tried from time to time." The same can be said for behavior as a criterion measure - no one pretends that it is perfect or all-wise, but, arguably, no better criterion exists.

*Empirical Research*

Few personality studies have actually obtained behavioral measures of traits. This is presumably because of the difficulties described above in obtaining objective, reliable measures of behavior. Furthermore, many of the studies that purport to measure behavior actually obtain self-reports of behavior. While this may satisfy the needs of certain researchers, it is almost useless for evaluating the accuracy of self- and informant-ratings of behavior. A good accuracy criterion should not be based on reports by the same people who provided the ratings being evaluated for accuracy. Thus, studies using self-report methods such as experience sampling or retrospective recall of behavior are not actually obtaining behavioral criteria in the strict sense.

Excluding those studies that rely on self- or informant-reports of behavior, most studies that measure behavior do so in a laboratory setting. Many such studies have come out of Funder's research group. For example, Kolar, Funder, and Colvin's (1996) measured people's behavior during a laboratory interaction with an opposite-sex stranger. Spain, Eaton, and Funder (2000) used a similar paradigm to measure extraverted behaviors in the laboratory. Since then, Furr, Funder, and Colvin (2000) have developed a behavioral rating system to code interactions in the laboratory. Borkenau and his research group have also made extensive use of laboratory-based measures of behavior as indicators of personality and intelligence (e.g., Borkenau & Liebler, 1993; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004). Another example is Gosling's study of behavior in group negotiation tasks (Gosling, John, Craik, & Robins, 1998). In all of these studies, great care was taken to obtain reliable, objective measures of behavior and to relate the behaviors to personality traits in theoretically meaningful ways. These studies should serve as models to personality researchers interested in collecting laboratory-based behavioral measures.

25

Even fewer studies have obtained naturalistic measures of behavior. The obstacles to collecting such data are clear. Behavioral measures are often quite obtrusive, which causes troubles for both the experimenter and the participant. One solution to this problem is to examine behavioral residue rather than direct behavior. For example, we can examine the contents of people's living spaces to see what kinds of objects they own and what traces of their behavior they have left behind (e.g., Gosling, Ko, Mannarelli, & Morris, 2002; Rentfrow & Gosling, 2003; Vazire & Gosling, 2004). Another solution is to use new technology to unobtrusively track people's everyday behavior. One example of such a method is the Electronically Activated Recorder (EAR; Mehl et al., 2001). This device allows researchers to record the sounds of people's lives by simply asking participants to wear a small digital recorder (that fits in a pocket) and a small lapel microphone. More details about this method are provided in Chapter 3.

The use of new technology and creative methods has allowed personality researchers to obtain objective, reliable, and valid measures of behavior. These developments have helped counter many of the limitations of behavioral research in personality, thus making behavioral criterion measures an appealing option for accuracy researchers.

**OVERVIEW OF THE STUDIES**

This dissertation has three main parts. In Chapter 2, I examine the differences among self-, friend-, partner-, parent-, and sibling-ratings of personality. Specifically, I compare how perceptions of personality differ across perspectives. Which perspective is the most harsh? The most lenient? Which perspectives agree most with each other? How much does each outside-perspective agree with the self's perspective? Are there some traits on which the perspectives agree more than others? What are the advantages and disadvantages of combining perspectives? What information is lost? And, finally, I carry

out a preliminary examination of the accuracy of the various perspectives using the behavioral criteria available to me.

Chapter 3 focuses exclusively on accuracy and compares the accuracy of self-perceptions to those by close others (friends, family, etc.). To lay the groundwork for future studies of accuracy, I assess a broad range of everyday behaviors and examine how well the self and others can predict how much the target performs these behaviors. What can people predict about their own behavior? What can others predict better? What is gained or lost by combining the two perspectives? I then go on to examine potential explanations for the patterns of accuracy in self- and other-ratings. What kinds of behaviors tend to be better predicted by the self? By others? I examine three potential moderators: the observability of the behavior, the desirability of the behavior, and the intentionality of the behavior.

Finally, in Chapter 4 I move away from predictions of behavior and attempt to evaluate the accuracy of self- and other-perceptions of personality. To do this I compare self-, friend-, and stranger-ratings of personality to objective, behavioral criteria for a handful of personality traits (e.g., anxiety, assertiveness, intelligence). The goal of this chapter is to determine what aspects of personality each perspective knows. The traits examined include many of the most important predictors of life outcomes in the domains of love, work, and health. By examining the accuracy of each perspective on a broad range of traits, I hope to expand our understanding of what people know about themselves and each other.

# Chapter 2: Five Perspectives on Personality

**OVERVIEW**

The purpose of Chapter 2 is to catalogue the differences among the self and close others' perceptions of personality. Because surprisingly little is known about how self-perceptions differ from other's perceptions, this exploratory study is needed to identify these differences. To make this study as comprehensive as possible, I examined five perspectives (self, friends, parents, romantic partners, and siblings) and nine traits. This study compares the perspectives in terms of three questions: positivity, self-other agreement, and, where possible, accuracy.

**BACKGROUND**

"Reality has always had too many heads."

-Bob Dylan

People are often seen differently by their different social groups. For example, an accountant may be seen as responsible and calm by his coworkers but his family may see him as absent-minded and emotional. Each group has a somewhat different perspective on the target, and each perspective represents a different version of reality. However, surprisingly little is known about whether there are systematic differences among these perspectives. How are people seen by their friends? Their families? Their romantic partners? How do these perspectives compare with the self's perspective?

Our friends, families, and romantic partners play an important role in our lives. What each of these groups thinks of us has an important causal force in shaping our interactions with them, the quality and outcome of these relationships (e.g., marital satisfaction and divorce)I propose that each perspective can be assessed along three

dimensions: (1) How positive are the perceptions? (Positivity), (2) How much do the perceptions agree with the targets' self-views? (Self-other agreement), and (3) How accurate are the perceptions? (Accuracy). This chapter addresses these three questions with respect to five perspectives: friends, parents, romantic partners, siblings, and, where applicable, the self. Throughout this chapter, I refer to these as "the five perspectives" or, where applicable, "the four perspectives" (referring to all of the non-self perspectives). The main thesis of this chapter is that the different perspectives have different patterns of positivity, self-other agreement, and accuracy. That is, the self, friends, parents, romantic partners, and siblings each have unique ways of perceiving the same targets, and these differences can be captured by levels of positivity, self-other agreement, and accuracy. For example, friends may be particularly accurate in their perceptions of targets' intelligence, whereas romantic partners may be particularly accurate in their perceptions of targets' depression. Such differences, which have until now been overlooked, would have important implications for understanding the interpersonal perception process, the unique perspective of the self, and for improving assessment in personality research, clinical diagnosis, and organizational settings.

The three questions I examine rest on the assumption that the five perspectives differ from one other. It should be noted, however, that I do not assume that the perspectives are completely different, or that because there are differences, they are necessarily wrong. It is well-established that even people who know the target in entirely different contexts (e.g., hometown friends versus college friends) agree substantially with each other about the target's personality (Funder, Kolar, & Blackman, 1995), and that their perceptions are to a large extent valid (e.g., Funder & Colvin, 1988). However, what has not been examined is the difference among the perspectives. I suspect that just as the overlap among the perspectives is meaningful (i.e., it reflects a person's shared

reputation), the differences among them may also be important for two reasons. First, understanding the differences among the perspectives will allow us to learn about the interpersonal perception process. For example, what leads a person to be seen differently by their friends than they are seen by their family? Second, understanding the differences among the perspectives will allow us to capitalize on these differences, instead of treating them like random error. Once we know what each perspective knows about a person, we will be able to determine who we should ask if we want to know about a specific trait.

Although virtually nothing is known about the differences among each of the perspectives, several theories, as well as research on self-other differences more generally, speak to the three questions and support my thesis that the perspectives differ in meaningful ways. In the sections that follow, I summarize the empirical and theoretical work related to these three questions based on the self-other literature, and, wherever possible, make specific predictions for each question. The questions and predictions are summarized in Table 1.

Table 1: Research Questions, Predictions, and Findings

| Research Questions | Predictions | Findings |
|---|---|---|
| 1. Which perspective is the most positive? | 2.1 Self and close others will be equally positive. | **Not supported.** Self less positive than close others. |
| | 2.2 Narcissism associated with greater self-other positivity discrepancy. | **Supported.** Narcissists see themselves more positively than close others see them. |
| | 2.3 Positivity will vary across traits. | **Supported.** Ratings more positive on more evaluative traits. |
| | 2.4 Positivity will vary across perspectives. | **Supported.** Parents most positive. Romantic partners selectively positive. |
| 2. Which perspective agrees most with the self? | 2.5 Self-other agreement will vary across traits. | **Supported**. Self-other agreement highest for extraversion, lowest for attractiveness, intelligence. |
| | 2.6 Self-other agreement will vary across perspectives. | **Tentatively supported.** E.g., parents marginally lower than other perspectives for emotional stability. |
| 3. Which perspective is most accurate? | 2.7 Accuracy will vary across traits. | **Supported.** |
| | 2.8 Accuracy will vary across perspectives. | **Supported.** |

**Positivity**

Do friends, parents, siblings, and romantic partners all see the targets equally positively? How does the positivity of the close others' ratings compare to the positivity of the targets' self-views? Do people see themselves more positively than anyone else sees them? Most research has ignored the question of how positively people are viewed by others, focusing exclusively on the positivity of self-views. One notable exception is the research showing that strangers rate targets more harshly than do acquaintances (John &

Robins, 1994; Kwan et al., 2004), but no research has directly examined the relative positivity of ratings by the self, friends, parents, siblings, and romantic partners specifically. Nevertheless, the two major self-enhancement theories and their associated lines of research speak indirectly to this question.

### *Positive illusions theory*

Positive illusions theory (Taylor & Brown, 1988) contends that most well-adjusted people have a positivity bias when it comes to their self-perceptions. One assumption of this line of research is that people see themselves more positively than others see them. Consistent with the positive illusions theory, empirical studies have demonstrated that most people see themselves more positively than strangers see them. However, there are large individual differences in the magnitude of this "bias" – my previous research has shown that only about 60% of people rate themselves more positively than strangers rate them (Vazire & Gosling, 2003). Furthermore, it is not clear that this finding is a result of self-enhancement. As others have pointed out, it is just as plausible that strangers have a stranger-harshness bias, and that their ratings are overly negative.

### *Realism theory*

Realism theory contends that most people's self-perceptions are realistic, and that overly positive self-views are associated with maladjustment in general (Colvin, Block, & Funder, 1995), and narcissism in particular (Paulhus & John, 1998; Robins & Beer, 2001). Consistent with realism theory, there is some preliminary evidence that the

positivity bias described above emerges only when comparing self-perceptions to strangers' ratings.

*Predictions*

The first aim of this chapter was to compare the positivity of the five perspectives. To provide a rigorous test of the two self-enhancement theories, I examined the degree of difference in positivity between self-views and close others' views. Consistent with realism theory, I predicted that there would be little or no difference in positivity between self-views and closer others' views (P2.1). In contrast, positive illusions theory would predict that self-views should be more positive than close others' views. To further test the two theories, I also examined individual differences in positivity discrepancies between self- and other-ratings, and predicted that, consistent with realism theory, narcissism would predict a greater discrepancy, with narcissists holding more positive views of themselves than others hold of them (P2.2).

I also predicted that positivity would vary across traits. Although no research to date has examined the properties of traits that predict positivity of self and others' ratings, I suspected that positivity would be greater for more evaluative and global traits (e.g., intelligence, attractiveness) than for more neutral and specific traits (e.g., openness to experience, conscientiousness). This prediction (P2.3) is based in part on positive illusions theory, which has mostly focused on global, evaluative traits such as intelligence.

However, these two theories focus exclusively on the positivity of self-views, and do not speak to the issue of positivity in perceptions of close others. The studies presented here provide the first empirical examination of this question. Based on the assumption that the perspectives differ, I predicted that the positivity of ratings would

vary across perspectives (P2.4). However, because of the dearth of research or theory on this topic, I did not make specific predictions about which perspective would be most positive for each trait.

**Self-Other Agreement**

How similar are people's self-views to the perceptions that their close others have of them? Previous research has found that self-other agreement correlations tends to fall in the .40 to .60 range for well-acquainted others, and several moderators of self-other agreement have been identified. Namely, self-other agreement is higher for more observable traits (Funder & Dobroth, 1987; Gosling et al., 1998; John & Robins, 1993) and lower for more evaluative traits (John & Robins, 1993). However, very little research has examined how self-other agreement is affected by the nature of the relationship between the self and the other. One exception is the literature on the effect of acquaintance on the self-other agreement (Kenny, 1994). The findings from this literature suggest that close others' ratings agree more with self-ratings than do strangers' ratings.

Another important exception is a study examining the interpersonal perception process in families (Branje et al., 2003). This study suggests that the nature of the relationship between perceiver and target may affect levels of self-other agreement, yet this has never been examined across different social groups outside of the family. For example, do people's self-views more closely resemble how they are seen by their friends or how they are seen by their parents?

Although virtually no empirical research has compared levels of self-other agreement among ratings by friends, parents, siblings, and romantic partners, three existing theories speak to this question, and provide support for the idea that the nature of the relationship between the self and the perceiver may have affect levels of self-other agreement.

*Self-verification*

One theory that provides a framework for making predictions about self-other agreement is self-verification theory. According to self-verification theory (Swann & Read, 1981), self-other agreement is the result, in part, of efforts by the target to bring the other person's perception of them in line with their target's own self-perceptions. According to this theory, people seek out others who verify their self-views, and people also manipulate others' views of them to achieve self-verification.

Thus, according to self-verification theory, self-other agreement should be stronger when the target is motivated and able to influence the other's views of the target. However, research has never examined which kinds of relationships are more likely to be the target of, or are more susceptible to, self-verification efforts. Thus, on the basis of self-verification theory, it is difficult to predict which perspective (friends, parents, siblings, or romantic partners) will agree most with the self.

*Selection, evocation, manipulation*

As Buss (1987) has argued, people interact with their environments to affect the way the see themselves and the ways others see them. First, people choose to enter in or avoid situations based on their perceptions of their personality and abilities. This process brings about the accuracy of self-views because self-views act as a self-fulfilling prophecy. It also constrains the behaviors and situations that others are able to observe, so this process may contribute to self-other agreement.

Second, people may change others' self-perceptions by virtue of the fact that they evoke certain behaviors in others. This may be conscious or unconscious. For example, if a woman sees her husband as nagging, she may evoke nagging behavior from him by ignoring his first, second, and third requests that she pick up after herself.

Finally, perceivers may bring targets to see themselves more like the perceivers see them. This may be an intentional manipulation, such as when a person convinces his friend to see himself as assertive. It can also be unintentional. For example, children's self-perceptions may be strongly influenced by their parents' perceptions of them, without parents intending to have this effect.

### *Weighted Average Model*

Kenny's Weighted Average Model (WAM; Kenny, 1994) also provides a framework for thinking about self-other agreement. According to this model, self-other agreement will be high to the extent that perceivers and targets share the same information, communicate with each other about their perceptions, and share the same interpretations of behavior.

The second aim of this chapter is to compare self-other agreement in ratings from friends, parents, siblings, and romantic partners for a broad range of traits. I made several predictions about self-other agreement. First, based on the research presented above, I predicted that self-other agreement would vary across traits (P2.5). Specifically, I predicted that self-other agreement would be higher for more observable and neutral traits. Second, based on the two theories presented here, I predicted that self-other agreement would be different across perspectives (P2.6). Once again, because of the dearth of previous research on this question, I did not make specific predictions about which perspective would agree most with the self for each trait.

### Accuracy

Perhaps the most important question about interpersonal perception is: how accurate are judgments by the self and others? Traditional accuracy research has conceptualized accuracy as a perceiver's agreement with the target's self-rating.

However, in the case of perceivers who are very well-acquainted with the target, it no longer makes sense to use the target's self-rating as the criterion for accuracy for most traits. For example, if people are seen as disagreeable by their close friends but they see themselves as agreeable, it is not obvious that the self is right and the close friends are wrong. For some personality traits the use of self-ratings as a criterion is still logical, even for high-acquaintance perceivers, because the traits are by definition self-views (e.g., depression, self-esteem). However, for most personality traits, an independent criterion is needed – one that is not based on self or close others' ratings.

Because of the difficulty of obtaining such "objective" measures, very little research has addressed this question empirically. However, two studies provide exceptionally rigorous tests. Kolar, Funder, and Colvin (1996) compared the ability of self-ratings and friends' ratings of behavior to predict participants' behavior in the laboratory. Their findings show that individual friends' ratings were slightly better than self-ratings at predicting laboratory behavior, and the aggregate of two friends' ratings greatly outperformed self-ratings. They also found that the self-ratings were especially inaccurate at predicting undesirable traits or behaviors.

Similarly, Spain, Eaton, and Funder (2000) compared the ability of self and close others' ratings on extraversion, neuroticism, and emotional experience to predict related outcomes. The criteria for extraversion and neuroticism ratings were behaviors coded from a laboratory activity, and the criterion for emotional experience ratings was self-reported emotional experience using a beeper methodology. Their findings indicate that the self was better for predicting emotional experience. However, this could be explained by method variance because the criterion for emotional experience ratings was a self-report measure. In addition, the criteria for all other ratings in both of these studies were

laboratory-based behaviors, thus limiting the ecological validity and generalizability of the findings.

Even less is known about the relative accuracy of ratings by the self, friends, parents, siblings, and romantic partners. Do siblings know something about their brother or sister that nobody else sees? Do romantic partners have particularly inaccurate views of their partner in certain domains? Although there is no empirical study addressing these questions, two theoretical models provide frameworks for examining this question.

### *Realistic Accuracy Model (RAM)*

Funder's RAM (1999) outlines the steps necessary to achieve accuracy in personality judgment, and can be applied to both self-perceptions and others' perceptions. The four steps of the model are: relevance, availability, detection, and utilization. The majority of the differences among the perspectives discussed here are likely to occur at the availability, detection, and utilization stages.

Availability refers to whether or not the perceiver (in this case the self, friend, parent, sibling, or romantic partner) has access to cues (e.g., behaviors) that are relevant for judging a particular trait. For example, if the only manifestations of neuroticism are very private behaviors such as crying or punching a wall, we would expect romantic partners to have greater availability of those behaviors than might friends, parents, or siblings, and this would lead to greater accuracy for romantic partners.

Detection refers to whether or not the perceiver detects the available cues. The five perspectives will likely differ in their success at this step due to differences in motivations to notice or ignore certain cues. For example someone who thinks of her brother as generous may not realize that he paid less than everyone else for a group dinner. In contrast, someone deciding whether or not to continue in a romantic

relationship (a position that college students who are dating are likely to find themselves in) may have a heightened awareness of their partner's characteristics.

Utilization refers to whether or not the perceiver correctly interprets the cues they have detected. The five perspectives are likely to differ in their success at this stage for three reasons: differences in motivation and reference groups. Motivational biases can affect how a perceiver interprets a cue. For example, a father who does not want to believe that his son is depressed could easily interpret his son's silence as manliness rather than a sign of depression. The perceiver's reference group will also affect how he or she interprets a cue. For example, a mother comparing her son to her own peer group may interpret his traveling to another state as a sign of extreme adventurousness where as her son's friends might not interpret his traveling as a particularly extreme behavior.

In short, differences among perspectives can occur at many different stages of the interpersonal perception process, resulting in different levels of accuracy. Among college student targets, parents and siblings may be at a particular disadvantage at the availability stage. However, the self may be at a greater disadvantage at the detection stage, because of biases or simply because they cannot directly observe their own behavior (due to the self's physical perspective). Although we do not yet know how the perspectives differ at each of these stages, the RAM model provides strong grounds for predicting that there are important differences in accuracy.

### *Pragmatic accuracy theory*

Pragmatic accuracy theory (Swann, 1984) extends the RAM model by taking into account how the perceiver's relationship with the target might influence the accuracy of their perceptions. According to pragmatic accuracy theory, people's perceptions of a target should be accurate to the extent that accuracy "facilitates the achievement of relationship-specific interaction goals." (Gill & Swann, 2004). Thus, we would expect the

accuracy of each perspective to differ across traits because the goals of friendships, romantic relationships, and family relationships are often very different. In addition, pragmatic accuracy theory stresses the importance of the context of the relationship between the perceiver and the target, and predicts that perceptions should be most accurate for those characteristics relevant to the context of their relationship. Thus, we would expect each perspective to have specialized knowledge of the target in different domains because each perspective knows the target in a different context.

The third aim of this chapter is to examine the relative accuracy of ratings by the self, friends, parents, siblings, and romantic partners. Both of the theories presented here provide many examples of why, how, and where the process of interpersonal perception might succeed or fail for each perspective. Thus, I predicted that accuracy would vary across traits (P2.7), and across perspectives (P2.8). Specifically, based on the research presented above, I predicted that accuracy would be particularly low for intelligence. However, because of the dearth of research on the different perspectives, I did not make any specific predictions about the relative accuracy of each perspective.

**Studies 2.1 and 2.2**

I conducted two studies to test my predictions. The studies were very similar to one another, but because the data were collected for different purposes, there are some small differences. Both studies included self-ratings as well as informant ratings by friends, parents, and romantic partners, and Study 2.2 also included ratings by siblings. Not all perspectives were available for all targets (e.g., some participants did not have a romantic partner), but the sample sizes were large enough to permit examinations of all of my research questions. In choosing which traits to examine, I used two criteria: breadth, and psychological importance. Thus, I assessed the five factors of the Big Five (see John & Srivastava, 1999; McCrae & Costa, 1999), allowing me to capture a broad

range of personality traits and at the same time make this research comparable with the substantial body of research using the Big Five. I also examined four other traits that are extremely important to people in choosing friends and mating partners (Botwin, Buss, & Shackelford, 1997; Buunk, Dijkstra, Fetchenhauer, & Kenrick, 2002; Kenrick, Groth, Trost, & Sadalla, 1993; Sprecher & Regan, 2002): intelligence, physical attractiveness, self-esteem, and depression. Although the traits were selected for their breadth an importance in everyday life, they also vary in evaluativeness, with intelligence and attractiveness being particularly evaluative, and in observability , with extraversion and physical attractiveness being particularly observable.

To compare the positivity of the five perspectives, I simply compared the mean level of ratings from each perspective for each trait. To compare self-other agreement among the perspectives, I computed self-other agreement correlations for each perspective (except the self) for each trait. Finally, to examine the accuracy of each perspective, I correlated the ratings from each perspective with a criterion measure of what the targets were actually like. For the Big Five, no clear, objective criterion has been developed. For intelligence, an IQ test scores served as the criterion. For attractiveness, strangers' ratings of the targets' overall physical attractiveness based on standardized photographs taken in the laboratory served as the criterion. Finally, for self-esteem and depression, self-ratings were used as the criterion because those traits are by definition self-views, and accuracy was only examined for the four close-other perspectives. Because I was only able to examine accuracy for these four traits, the accuracy analyses presented here should be considered a preliminary glimpse at the relative accuracy of these perspectives for a few important traits, intended to serve as the groundwork for more comprehensive examinations in future studies.

**METHOD**

**Study 2.1**

The data from Study 2.1 presented here are a subset of a larger study. I provide a description of the participants, as well as the measures and procedures relevant to the analyses presented in this chapter.

*Participants*

Participants were 80 undergraduate students recruited mainly from Introductory Psychology courses and flyers in the Psychology department. The sample was 54% female, 65% White, 21% Asian, 11% Latino, and 3% of another ethnicity. The vast majority of the participants (73%) were 18 years old, although their ages ranged from 18 to 24 (M = 18.7, SD = 1.4). Participants received $50 in compensation for complete participation in the large-scale study (all participants completed the entire study).

*Informants*

Each participant was asked to nominate three people who knew them well to provide ratings of their personality. Participants were asked to nominate one friend, one parent, and one romantic partner if possible. If participants could not provide any one of the three kinds of informants, they were told to nominate someone who knows them well. Participants were told that the informants' ratings would be kept completely confidential, and that they themselves would never see their informants' ratings.

Two months after the end of the study, 76% of informants had completed the ratings, resulting in a total of 182 informant ratings, of which 87 were from friends, 55 from parents, and 21 from romantic partners. The remaining informant reports were from siblings (12), ex-romantic partners (2), and one cousin, one grandmother, one great-

grandmother, one ex-friend, and one informant who did not indicate his or her relationship to the target.

*Measures*

Participants and informants completed a battery of measures, which included the Big Five Inventory (BFI; John & Srivastava, 1999) and single-item measures of intelligence, physical attractiveness, and self-esteem. Informants also rated the targets on a single-item measure of depression, and self-ratings of depression were obtained from both the BFI item "is depressed, blue" which was used in the positivity analyses to make the self-rating mean comparable with the single-item informant rating means, and the depression facet of neuroticism on the NEO Personality Inventory – Revised (NEO PI-R; Costa & McCrae, 1985) which was used in the accuracy analyses because it is more reliable than the BFI item. All ratings were made on a 7-point Likert-type scale. Reliabilities for the self-ratings and informant ratings on the BFI were .89 and .91 respectively, for extraversion, .79 and .92 for agreeableness, .76 and .89 for conscientiousness, .81 and .88 for emotional stability, and .85 and .90 for openness. The reliability of the self-ratings on the depression facet of the NEO PI-R was .85.

Participants also completed the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988). The reliability of NPI scores was .83.

*Procedure*

All self-report measures were obtained on the first day of the three-week study. After consenting to participate, participants completed the informant-nomination questionnaire, followed by a battery of self-report questionnaires, including all the measures described above. The self-report questionnaires were administered online, on one of the laboratory computers. Participants were seated in a private room with the door

closed, and were told to complete the questionnaires online, taking as many breaks as they desired. None of the data collected after these procedures were used in the analyses presented here.

As recommended by Vazire (in press-a), informant-ratings were collected via the Internet. Informants were contacted by e-mail and asked to complete an online questionnaire about how they see the target participants' personality. Informants received a link and unique identifying number in the email. Informants who did not complete the ratings were sent reminder emails after two weeks, four weeks, and six weeks. Participants were compensated at the end of the three weeks, regardless of whether the informants had completed their ratings. Informants were not compensated for their cooperation. Previous research has shown that Internet questionnaires are a valid method for collecting personality ratings (Gosling, Vazire, Srivastava, & John, 2004).

**Study 2.2**

The data from Study 2.2 presented here are also a subset of a larger study. Like Study 2.1, this study also involved multiple sessions. The analyses in this chapter include data from both sessions of the study. The two sessions were approximately two months apart, and 97% of participants completed both sessions.

*Participants*

Participants were 160 undergraduate students enrolled in Introductory Psychology at the University of Texas at Austin. The sample was 53% female, 56% White, 23% Asian, 12% Latino, 3% Black, and 3% of another ethnicity. The sample was largely comprised of 18 and 19 year olds (84%) but the ages of participants ranged from 17 to 40 (M = 18.7, SD = 2.0). Participants received partial course credit in return for their participation. Of the original 160 participants, 155 returned for the second session.

Each participant was asked to nominate three people who knew them well to provide ratings of their personality. In this study, participants were simply told to nominate people who knew them well, and were not asked specifically for one of each kind of informant. Participants were told that the informants' ratings would be kept completely confidential, and that they themselves would never see their informants' ratings. Because self-reports were collected during the second session, five participants did not complete the second session. Therefore, my analyses only used informant reports of those 155 participants who completed both sessions.

Two months after the end of the study, 82% of informants had completed the ratings, resulting in a total of 381 informant ratings, of which 217 were from friends, 63 from parents, 44 from siblings, and 41 from romantic partners. The remaining informant reports were from ex-romantic partners (4), cousins (6), and one aunt, one uncle, one son, and three informants who did not indicate their relationship to the target.

*Measures*

Participants and informants completed a battery of measures, which included the BFI and single-item measures of intelligence, physical attractiveness, self-esteem, and depression. All ratings were made on a 7-point Likert-type scale. Reliabilities for the self-ratings and informant ratings on the BFI were .86 and .72, respectively, for extraversion, .82 and .80 for agreeableness, .78 and .75 for conscientiousness, .84 and .80 for emotional stability, and .76 and .76 for openness. Participants also completed the NPI, which had a reliability of .87. Participants' IQ was measured using the Wonderlic Personnel test (Wonderlic, 1983), a 12-minute test measuring both verbal and non-verbal IQ. Finally, participants' attractiveness was measured by having 12 judges rate the

targets' physical attractiveness on the basis of a photograph on a single item using a 7-point Likert-type scale. The reliability of the judges' ratings was .87.

### *Procedure*

All self-report measures were obtained during the second of two sessions of the study. Other than this, the procedure for the self-report portion of the study was identical to the procedure in Study 2.1. In addition, criterion measures were obtained for intelligence and physical attractiveness. Participants completed the Wonderlic IQ test using traditional paper-and-pencil methods, while sitting in a room by themselves with the door closed. The test was administered during the first session, approximately two months before the self-reports of intelligence were collected. Participants did not receive feedback on their performance on the IQ test. Participants' total score on the IQ test served as the criterion for self and informant ratings of intelligence.

The photographs of participants used for judging physical attractiveness were taken at both sessions and shown to two groups of six judges (each group of judges only saw one session). The photographs were taken using the same standard procedure for all participants. Participants were asked to stand against a white wall in a bare room, containing only a camera on a tripod. The location of the camera and the participant were fixed so that the bottom of the frame was just below the participant's feet, thus ensuring that the entire body would be captured in the photograph. During the first session, participants did not know before coming to the experiment that they would be photographed. During the second session, they had been told ahead of time that they would be photographed again. The photographs were then shown to the judges on a photo CD, and the judges made their ratings on a website. The judges' ratings were aggregated across both sessions and this average was used as the criterion for self and informant ratings of physical attractiveness.

46

The procedure for collecting informant reports via the Internet was identical to the procedure described in Study 2.1. Participants were compensated at the end of the second session, regardless of whether the informants had completed their ratings. Informants were not compensated for their cooperation.

**RESULTS**

**Positivity**

To compare the positivity of the five perspectives' ratings, I computed mean scores for each perspective on each of the nine traits. These means for both studies are presented in Figure 1. The most striking finding is that, contrary to my prediction (P2.1) and contrary to positive illusions theory, the perspective of the self was the least positive for almost all traits, and this finding replicated in both studies. For Study 2.2, paired samples t-tests revealed that self-ratings were significantly less positive than parents' ratings for every trait except depression (all $p$'s < 2.26, all $p$'s < .05). In addition, self-ratings were significantly less positive than all other perspectives' ratings for extraversion, agreeableness, intelligence, and physical attractiveness (all $t$'s > 2.12, all $p$'s < .05). For conscientiousness, self-ratings were significantly less positive than friends' ratings ($t (1, 103) = 4.93$, $p < .01$), parents' ratings ($t (1, 43) = 7.64$, $p < .01$), and romantic partners' ratings ($t (1, 37) = 3.25$, $p < .01$). For emotional stability, openness, and self-esteem, self-ratings were not significantly different from any other perspectives' ratings (except for parents' ratings, as mentioned above). Finally, for depression, self-ratings were not significantly different from any other perspectives' ratings.

Figure 1: Positivity of Ratings by Self, Friends, Parents, Partners, and Siblings
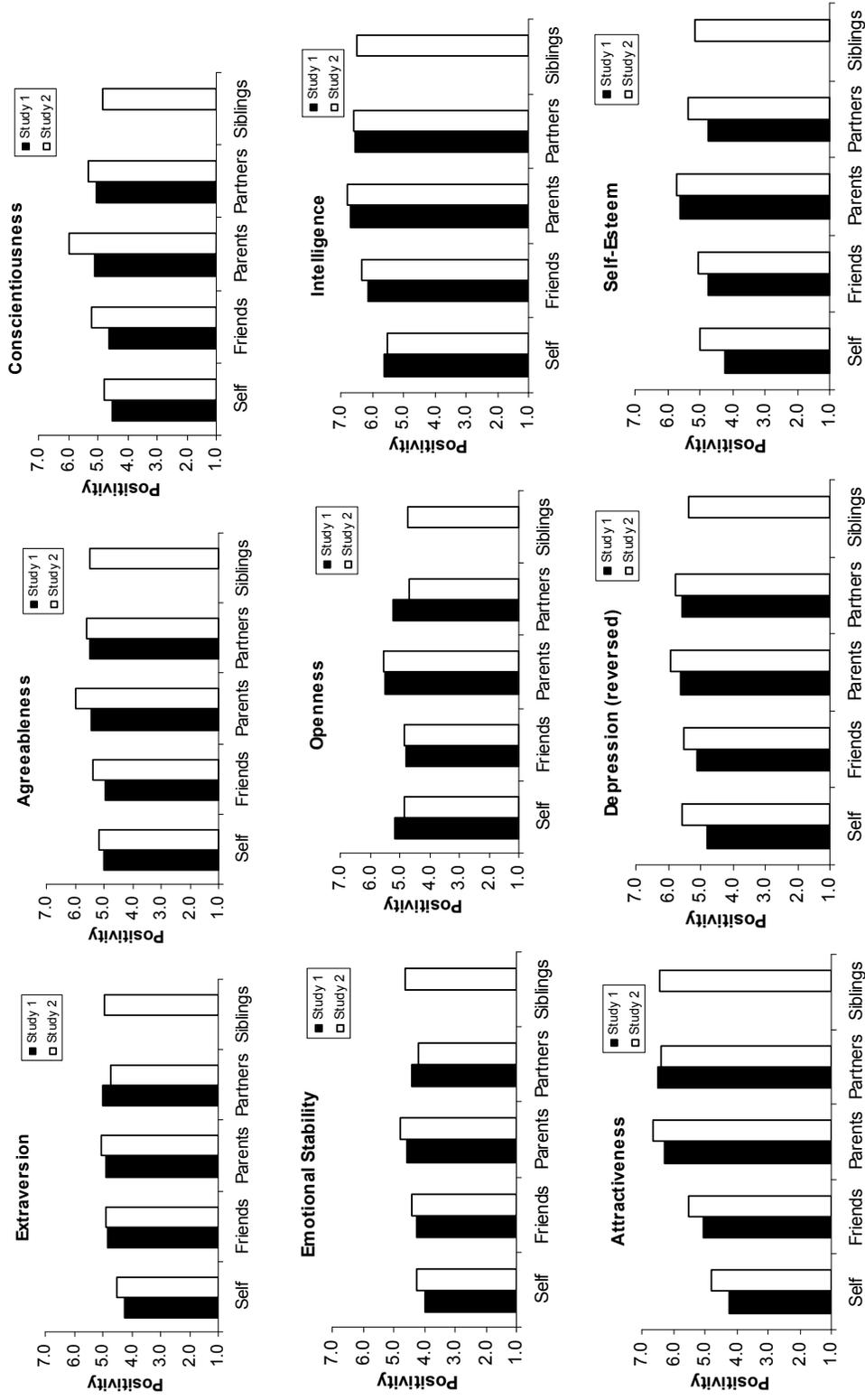
To test my prediction that narcissism would be associated with greater self-other discrepancy (P2.2), I computed average-other scores for each trait by averaging across the three informants' ratings for each participant. I then computed self-other difference scores for each trait by subtracting the average-other ratings from the self-ratings. Finally, I correlated the difference scores for each trait with the targets' NPI scores. Consistent with my prediction (P2.2) and with realism theory, narcissism was positively and significantly correlated with self-other difference scores for conscientiousness ($r = .21$, $p < .05$), emotional stability ($r = .18$, $p < .05$), intelligence ($r = .30$, $p < .01$) and attractiveness ($r = .36$, $p < .01$). That is, narcissists were more likely than non-narcissists to see themselves more positively on these traits than their close others saw them.

Consistent with my prediction (P2.3), positivity did vary across traits. Specifically, as I predicted, ratings were most positive for the more global and evaluative traits: intelligence and attractiveness.

Consistent with my prediction (P2.4), positivity also varied across perspectives. Most striking was the finding that parents' ratings were the most positive for almost every trait across both studies. In addition, more complex patterns also emerged from the findings. Specifically, romantic partner's ratings were generally very positive (especially for intelligence and attractiveness), but were less positive than most other perspectives for affective traits (i.e., emotional stability, self-esteem, and depression), particularly in Study 2.2. In addition, friends were particularly harsh (relative to the other non-self perspectives) in their ratings of intelligence and attractiveness, and this finding replicated across both studies.

In summary, three of my four predictions for positivity were supported. Self-ratings were less positive than others' ratings (contrary to P2.1), and self-other discrepancy was related to narcissism such that narcissists were more like to rate

themselves more positively than their close others rated them (consistent with P2.2). In addition, positivity did vary across traits (consistent with P2.3) and across perspectives (consistent with P2.4).

**Self-Other Agreement**

Table 2 shows the levels of self-other agreement across seven traits for both studies. Self-esteem and depression were omitted from these analyses because the self-ratings were used as the criterion for accuracy, and so self-other agreement correlations are identical to the accuracy correlations reported in the next section. The correlations in Table 2 represent the correlation between each perspective's ratings of the targets on a trait, and the targets' self-ratings on the same trait.

Notice that the mean levels of overall self-other agreement (across all traits) were not very different among the four perspectives. The last row of Table 2 shows that the average self-other agreement correlations ranged from .30 to .46 for the four perspectives. These values are consistent with previous research showing that levels of self-other agreement for well-acquainted informants typically range from .40 to .60 for the Big Five personality traits (Watson, Hubbard, & Wiese, 2000).

Consistent with my prediction (P2.5), self-other agreement did vary across traits. Recall that, based on previous research, I predicted that self-other agreement would be higher for more observable traits. These findings provide partial support for this prediction. Both extraversion and physical attractiveness are relatively observable traits, yet self-other agreement was very high for extraversion and relatively low for attractiveness across both studies and almost all perspectives.

50

Table 2: Self-Other Agreement for Each Perspective

| | Study 2.1 | | |
|---|---|---|---|
| | Friends | Parents | Partners |
| | N = 66 [a] | N = 51 [b] | N = 21 |
| Extraversion | .56** | .61** | .65** |
| Agreeableness | .37** a | .65** b | .73** b |
| Conscientiousness | .58** | .48** | .05 |
| Emotional Stability | .45** | .38** | .45* |
| Openness | .24* a | .43** a | .68** b |
| Intelligence | .08 ab | .27* a | -.03 b |
| Attractiveness | .27* | .31* | .29 |
| Average | .38 | .46 | .44 |

| | Study 2.2 | | | |
|---|---|---|---|---|
| | Friends | Parents | Partners | Siblings |
| | N = 106 [c] | N = 46 [d] | N = 41 | N = 33 [e] |
| Extraversion | .67** a | .57** ab | .76** a | .42** b |
| Agreeableness | .53** | .34* | .28* | .49** |
| Conscientiousness | .26** | .20 | .30* | .33* |
| Emotional Stability | .49** a | .28* ab | .53** a | .19 b |
| Openness | .48** | .42** | .17 | .28 |
| Intelligence | .31** | .33* | .27* | .34* |
| Attractiveness | .25** | .18 | .30* | -.12 |
| Average | .42 | .33 | .44 | .30 |

*Note.* All values are Pearson's r's. a 87 friends rated 66 targets. b 55 parents rated 51 targets. c 217 friends rated 106 targets. d 63 parents rated 46 targets. e 44 siblings rated 33 targets. Correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed. * $p < .05$, two-tailed; ** $p < .01$, two-tailed.

Although the perspectives did not differ in their overall levels of self-other agreement across traits, the findings do show that each perspective had unique patterns of self-other agreement across traits (P2.6). For example, although the partners had the

highest levels of self-other agreement on extraversion across both studies, they also had the lowest levels of self-other agreement on intelligence across both studies. Another pattern that was apparent across the two studies was the low levels of self-other agreement for parents on emotional stability. However, due to the small sample size for some of the perspectives, none of these patterns were statistically significant. Thus, although the findings provide preliminary support for my prediction (P2.6), more power is needed to detect the specific patterns of self-other agreement across traits and perspectives with greater certainty.

**Accuracy**

Table 3 shows the levels of accuracy for each perspective on the four traits for which an accuracy criterion was available. Recall that the criterion for intelligence ratings was the target's IQ test score, the criterion for attractiveness ratings was the target's aggregate rating of attractiveness from 12 judges who viewed a photograph of the target, and the criteria for both self-esteem and depression were the targets' self-ratings. The correlations in Table 3 represent the correlation between each perspective's ratings of the targets on a trait and the criterion for that trait.

Consistent with my prediction (P2.7), accuracy did vary across traits. Specifically, accuracy was lower for intelligence and attractiveness than for self-esteem and depression. This is consistent with previous research showing that self- and other-ratings of intelligence are not very accurate.

Table 3: Accuracy of Ratings from Each Perspective

| | Study 2.1 | | |
|---|---|---|---|
| | Friends N = 66 [a] | Parents N = 51 [b] | Partners N = 21 |
| Self-esteem | .27* | .16 | .32$^{\dagger}$ |
| Depression | .28* | .45* | .32$^{\dagger}$ |

| | Study 2.2 | | | | |
|---|---|---|---|---|---|
| | Self N = 155 | Friends N = 106 [c] | Parents N = 46 [d] | Partners N = 41 | Siblings N = 33 [e] |
| Intelligence | .22** | .18* | .16 | .11 | .21 |
| Attractiveness | .18* [a] | .34** [b] | .07 [a] | .12 [a] | .31* [ab] |
| Self-esteem | | .45** | .29* | .45** | .42** |
| Depression | | .28** [a] | .35** [a] | .66** [b] | .33* [a] |

*Note.* All values are Pearson's r's. [a] 87 friends rated 66 targets. [b] 55 parents rated 51 targets. [c] 217 friends rated 106 targets. [d] 63 parents rated 46 targets. [e] 44 siblings rated 33 targets. The criterion for ratings of intelligence was an IQ test score. The criterion for ratings of attractiveness was the aggregate of judges' ratings of attractiveness based on photographs of the targets. The criterion for self-esteem and depression was the targets' self-ratings. Within each study, correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed. $^{\dagger}$ p < .10, two-tailed; * p < .05, two-tailed; ** p < .01, two-tailed.

Consistent with my prediction (P2.8), accuracy also varied across perspectives. The pattern of accuracy across perspectives was very different for different traits. Specifically, for attractiveness, friends and siblings were the most accurate while parents, romantic partners, and the self were not accurate. In contrast, for ratings of self-esteem and depression, friends and romantic partners were the most accurate while parents were generally less accurate, and siblings fell somewhere in between.

**DISCUSSION**

**Positivity**

Three out of the four predictions about positivity were supported. The finding that self-ratings were less positive than the others' ratings on almost all traits, across both studies, was unexpected. Positive illusions theory would have predicted that the self-ratings would be more positive than the others' ratings. Indeed, even realism theory would have predicted that the self-ratings would be about as positive as others' ratings. Why did I find such a strong self-diminishment effect?

There are three possible explanations for the findings. First, it is possible that the participants in these studies were being falsely modest. That is, they actually hold more positive views of themselves than they reported. Although this is possible, there is no reason to believe that these participants were more concerned with modest self-presentation than were participants in previous studies of self-enhancement. Thus, these participants' self-reports were probably not less positive than the self-reports of other participants in other studies, even if there was a false-modesty effect. Following this logic, false modesty cannot account for the inconsistency between the present findings and positive illusions theory.

Another possible explanation is that people's self-views actually are not unrealistically positive. That is, perhaps the participants in the current sample were reporting honestly, and their close others' very positive ratings of them reflect the fact that their self-views are, if anything, overly negative. If this is the case, positive illusions theory and realism theory are both wrong, and people actually hold overly negative views of themselves. This is unlikely given the fairly convincing research showing that people do tend to like themselves, be optimistic about their own future, and see themselves more positively than they see anonymous others (Taylor & Brown, 1988).

54

A third, more plausible explanation is that enhancement is not limited to the self. That is, perhaps our friends, parents, romantic partners, and siblings actually hold more unrealistically positive views of us than we hold of ourselves. This explanation does not rule out the possibility that self-enhancement also occurs, it simply holds that we enhance our close others more than we enhance ourselves, or that we choose to be close to people who see us positively.

Although more research is needed to ascertain the reason for the self-diminishment effect found here, I suspect that other-enhancement is an important phenomenon that is largely responsible for this finding. If this is true, it has important implications for theories of the self. We know a little bit about why people idealize their romantic partners, but why do we see even our friends and family members more positively than they see themselves? Broader theories of the self and social relationships (e.g., evolutionary theory) may be able to illuminate this important issue, which has hitherto been overlooked.

Our finding that narcissism was associated with greater discrepancy between self-views and close others' views is consistent with my prediction and with previous research on realism theory. This finding provides further evidence that self-enhancement, at its extreme, is a reflection of an unhealthy, narcissistic personality. Because self-other discrepancies are a result of both positive self-ratings and negative ratings from others, it is not clear whether the causal direction of the effect in this study was from narcissism to self-other discrepancy or the other way around. That is, does narcissism lead to being seen negatively by others, or does being seen negatively by others lead to narcissism? Previous research on narcissism suggests that it is the former. Narcissists tend to be relatively maladjusted (Colvin, Block, & Funder, 1995), disliked by their peers (in the

long-run; Paulhus, 1998), and experience declines in self-esteem and academic engagement (Robins & Beer, 2001).

The finding that positivity was highest for the more global and evaluative traits (intelligence and attractiveness) sheds light on the inconsistencies between positive illusions theory and realism theory. Research on positive illusions theory has traditionally examined global, evaluative constructs such as intelligence and has found that self-ratings on these constructs tend to be extremely positive. In contrast, much of the research from the realism perspective has examined more specific traits, such as California Q-Sort items (Colvin, Block, & Funder, 1995), and has found that most people do not self-enhance on these items. The findings presented here suggest that the conflicting findings may be due to the different item content used by the two groups of researchers – people may rate themselves (and close others) especially positively on global, evaluative constructs.

Finally, the finding that positivity varied across perspectives supports the main thesis of this chapter, that there are meaningful, systematic differences among the five perspectives. Across both studies, parents were the most positive of any perspective for almost all traits, and romantic partners were selectively positive (i.e., positive on some traits but not others), demonstrating that, when it comes to positivity, all close others are not equal.

**Self-Other Agreement**

Previous theorizing on self-other agreement has proposed that it is moderated by trait observability. Consistent with this notion, I found that self-other agreement was highest for extraversion, a highly observable trait. However, I also found that self-other agreement was low for physical attractiveness, arguably the most observable trait in this study. This finding suggests that a trait's observability may be less important than a

56

trait's evaluativeness in predicting self-other agreement. Consistent with this idea, self-other agreement was also relatively low for intelligence, another highly evaluative trait.

I also predicted that self-other agreement would vary across perspectives. Although there was no overall mean difference in self-other agreement among the perspectives, there was preliminary evidence that patterns of self-other agreement across traits were different for each perspective. For example, self-other agreement on emotional stability was marginally lower for parents than for other perspectives. However, more research with larger groups of informants is needed to uncover the specific patterns with greater certainty.

Once these patterns are established, the next step will be to examine the processes underlying these differences. For example, do people engage in differential self-verification with their different social groups? It is possible that people bring their romantic partners to see them as they see themselves on attractiveness and emotional stability, but bring parents to see them as they see themselves on openness and intelligence. It is also possible that, consistent with the looking-glass self theory, self-other agreement results from close others influencing self-views. For example, perhaps people come to see their own intelligence through their parents' eyes, thus leading to higher self-parent agreement on intelligence. In short, understanding how self-other agreement varies across perspectives will lead to a better understanding of how people present themselves differently to their various social groups, and how one's relationship to a target influences our perception of that target.

**Accuracy**

One clear reflection of how the perspectives differ is their relative accuracy. Close others are often used as informants in assessing personality, diagnosing mental illness, and assessing job performance or potential. However, researchers have rarely examined

57

what kinds of informants provide the most accurate information. The present findings suggest that the answer to this question depends on the trait being assessed. No one perspective is omniscient, therefore, assessment accuracy would be improved if researchers exploited the strengths of each perspective. Because I could only assess accuracy for four traits, I can only draw limited conclusions about the domains of expertise for each perspective. For example, the present findings suggest that romantic partners may be especially attuned to levels of self-esteem and depression, whereas parents seem particularly unaware of these traits. Friends and siblings are particularly accurate at rating attractiveness while the self, parents, and romantic partners would not be good informants for this trait. These patterns once again support my main thesis that the different perspectives differ in meaningful and systematic ways.

Why do these patterns emerge? The RAM model of personality judgment would suggest that differences in availability, detection, or utilization may explain differences in accuracy among perspectives (Funder, 1999). For example, perhaps romantic partners have more opportunities to observe depression-related behaviors. In contrast, pragmatic accuracy theory would explain these differences in terms of how relevant and important each trait is for each perspective. For example, depression might be particularly relevant to romantic relationships, and less relevant to parent-child interactions. Future research should test these competing hypotheses by examining the processes underlying the perceptions from the self, friends, parents, romantic partners, and siblings.

All five perspectives were quite bad at rating intelligence. These low correlations may be due in part to a restriction of range in the ratings of intelligence, considering the extremely elevated means for intelligence presented in Table 2 However, the low levels of accuracy for intelligence are consistent with previous research, and probably reflect a genuine lack of awareness on the part of the raters. This finding and the other low

correlations in Table 3 present an interesting question for future research: why are people who know the targets so well (including the targets themselves) so bad at perceiving traits as fundamental and important as intelligence, attractiveness, and self-esteem? And, conversely, how do some perspectives come to know so much about these traits? Continuing to examine these questions will provide important insights into how personality judgment succeeds and fails.

The findings in chapter 2 are important for two reasons. First, theories of interpersonal perception (such as RAM; Funder, 1999) can be applied to the patterns of findings in order to shed light on how differences among the perspectives emerge. Such applications of theoretical models help refine our understanding of the interpersonal perception process more generally. Second, understanding how the perspectives differ can help us improve the quality of personality assessment by capitalizing on the each perspective's strengths. This information will benefit personality researchers, clinicians, and organizations interested in obtaining a more nuanced and accurate portrait of people.

In summary, the studies presented in chapter 2 revealed important differences among perceptions of the same people by themselves, their friends, their parents, their romantic partners, and their siblings. The perspectives differ in terms of the traits on which they enhance the target person, the traits on which they most agree with the target person's self-views, and the traits on which they hold the most accurate perceptions. Once again, Bob Dylan's lyrics are apt: "Half of the people can be part right all of the time, some of the people can be all right part of the time, but all the people can't be all right all the time."

# Chapter 3: The Accuracy of Self- and Other-predictions of Behavior

**OVERVIEW**

Chapter 2 examined the differences among different perspectives on a person, but only touched upon the question of which perspective is more accurate. The purpose of Chapter 3 is to examine the accuracy of self- and other-perceptions, and to develop a model for cataloguing and explaining differences in self- and other-knowledge.

**BACKGROUND**

"You can't see in and it's hard looking out."

-Bob Dylan

Who knows a person best? Are people their own best expert, or do others know something about them that they don't know about themselves? The major difference between the two perspectives is that, as Bob Dylan's quote implies, the self and the others each have a unique angle on the target's behavior. The self is on the inside trying to look out and observe their own behavior, while the others are on the outside and can only infer what is going on inside. As a result, each perspective is likely to have unique knowledge, and unique blind spots. According to this view, both perspectives should contribute uniquely to the accuracy of behavioral predictions. In this chapter, I compare the two perspectives (self vs. other) in terms of how well each can predict behavior.

Predicting behavior is of utmost importance in everyday life. We base our decisions of who to date, vote for, trust, work with, hire, and marry based on our assumptions of how they will behave in the future. In addition, predicting how people will behave is one of the main goals of research in personality and social psychology. In short, to really know someone is to be able to predict how they will behave. Are they the

kind of person who shows up on time? Are they likely to get in a fight? Do they spend most of their time with others or alone?

What is the best way to predict someone's behavior? Previous research has shown that self-predictions of behavior are not very accurate. People do not know how they will behave in the future. In fact, they are not even very aware of how they behaved in the past (Gosling et al., 1998). This poses a serious threat to research – and lay judgments – based on self-reports of behavior. Considering the difficulty of measuring actual behavior (and the impossibility of measuring future behavior), what can researchers do to improve the validity of predictions of behavior? How can we better predict what a person will do in the future?

To answer this question, we can look to the strategies people use in everyday life to make more informed predictions of one another. When trying to predict someone's behavior, we rarely rely exclusively on self-descriptions. This may be because we cannot ask the person themselves, because we do not trust that the person will be honest, or because we suspect that the person is simply unaware of his or her own behavior. In these cases, people often turn to others for a second (or first) opinion. For example, a woman once wrote in to an advice column saying that her mother's coworkers had told her that the man she is seeing is a womanizer. Dear Prudence, the advice columnist, wrote back: "It is Prudie's hunch that if you continue with this man, you will learn that your mother's 'informants' were correct." (Slate.com, February 10, 2005). As this example illustrates, we often turn to informants to help us make better predictions about how a person will behave. This chapter examines the validity of self and informant-reports of behavior. Do others' descriptions actually improve the accuracy of predictions? Can others tell us something that the self cannot?

I begin by presenting a model of self- and other-knowledge that serves as a framework for examining the accuracy of the two perspectives. Our model is meant to serve as a framework for building theories about the processes underlying the two perspectives. What factors influence the accuracy of self-perceptions and others' perceptions? What kinds of behaviors is the self most aware of, and what kinds of behaviors are others more aware of? To begin with, I propose three dimensions of behavior that may affect the accuracy of self and others' views: observability, desirability, and automaticity. I then present two studies that test these three potential moderators. The second study also presents exploratory analyses to identify the behaviors that self or others are more accurate on. The results from these studies permit a refinement of the model and develop more specific hypotheses about why and how self and other perceptions differ. In addition to developing a framework for self- and other-knowledge, the present research improves upon previous studies on the accuracy of self and others' predictions by using actual real-world behavior as the criterion for accuracy in Study 3.2.

**THE MODEL**

The basic tenet of my model is very simple: what people know about themselves and what others know about them are overlapping but not identical domains. Thus, there are some things about a person that only they themselves know, some things that only others know, some things that both the self and others know, and some things that nobody knows. Figure 2 illustrates this model.

Figure 2: Model of Self- and Other-Knowledge

Self-Knowledge

|  | Low | High |
|---|---|---|
| **High** | Only Others Accurate | Self & Others Accurate |
| **Low** | Nobody Accurate | Only Self Accurate |

Other-Knowledge
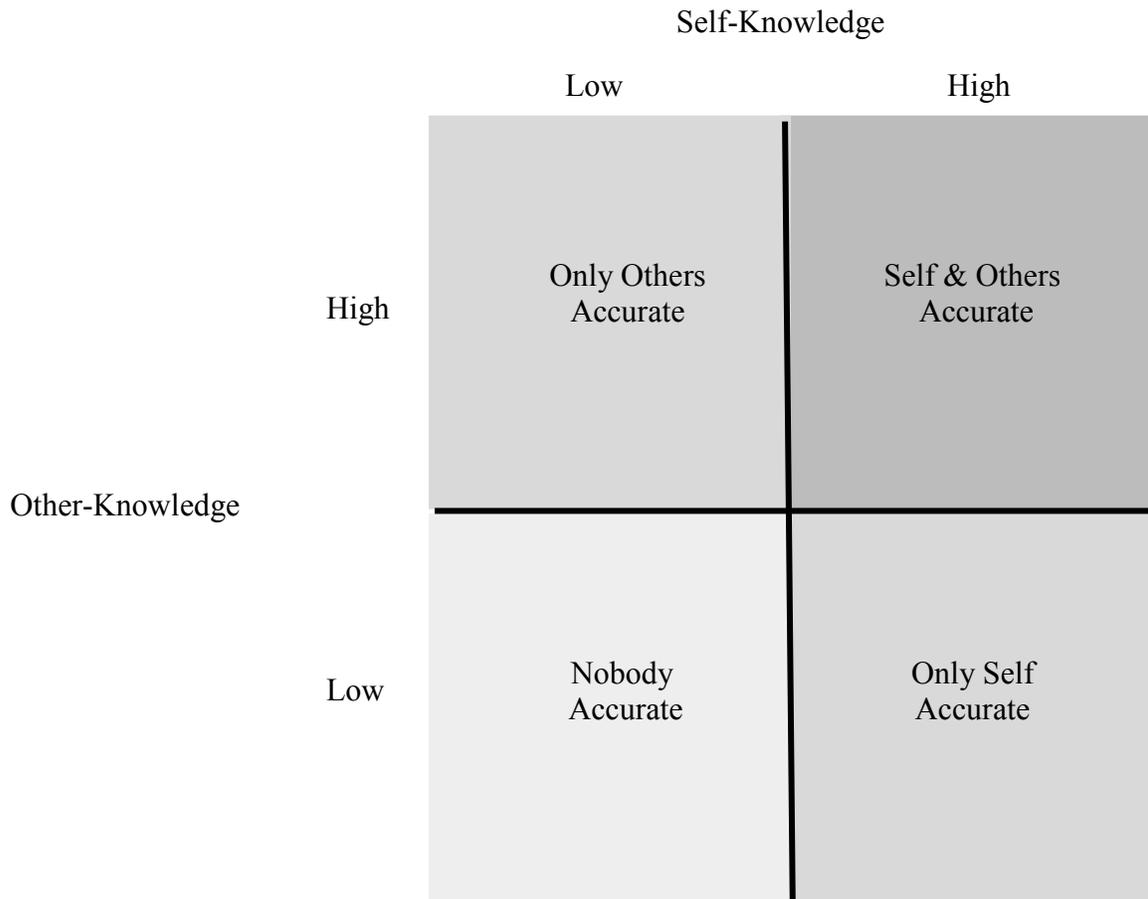
Although simple, this model is important for two reasons. First, it assumes that there are some things people don't know about themselves, an idea that, until recently, has not received much attention. In the last few decades, researchers have relied almost exclusively on self-reports when assessing personality, behavior, values, and affect, and have often assumed that self-views are a relatively accurate reflection of reality. Only recently have researchers acknowledged the limits of self-knowledge, and rediscovered the benefits of collecting other sources of information such as informant reports (e.g.,

Vazire, in press-a), implicit measures, experience sampling, and physiological measures. However, the idea that self-knowledge is limited, and indeed that others may know things about us that we don't know about ourselves, is still controversial. Thus, the first and most important prediction of my model is that the domains of self- and other-knowledge are overlapping but not identical. That is, there are some behaviors that only the self knows about (i.e., can report on accurately) and some that only others know about (P3.1).

Second, this model is the first step to identifying the underlying processes that lead to differences between the two perspectives. The ultimate purpose of this model is to identify the dimensions that distinguish what the self knows from what others know. In order to do this, researchers must first conduct exploratory research placing specific behaviors in the various parts of the model. The primary goal of this chapter is to place a broad range of everyday behaviors in the context of this model in order to get a better picture of the landscape of self- and other-knowledge.

This model closely resembles the Johari window, a model proposed by Luft and Ingham (1955). Luft and Ingham proposed this as a model for self-awareness, but the model has gotten little attention among researchers (though it is used occasionally in therapeutic settings). Although little is know about where various behaviors fall in my model, some research exists on the moderating effects of dimensions such as observability and evaluativeness (Cheek, 1982; Robins & John, 1997). I draw on this research and existing theories to propose three dimensions of behavior that may distinguish what the self knows from what others know: observability, desirability, and automaticity. As I explain in detail below, I predict that all three of these dimensions will moderate differences in self- and other-knowledge. Identifying the kinds of behaviors that only the self knows or only others know, and how each perspective comes to have this

unique knowledge, will help us predict behavior better and understand the nature of self- and other-perception.

**PRESENT STUDIES**

**Study 3.1**

The first study I present here examined people's lay theories about what the self and others know about a person. To examine lay theories, I asked people to rate a number of behaviors on two dimensions: how accurate people are at predicting this behavior in themselves, and how accurate people are at predicting this behavior in others they know well. The main purpose of this study was to paint a picture of people's lay beliefs about what behaviors fall into which quadrants of the model, and to determine whether lay theories are consistent with the idea that the self and others have overlapping but unique domains of knowledge. I also examined whether the patterns of self- and other-knowledge in lay theories were moderated by the observability, desirability, and automaticity of the behaviors. That is, I tested my hypotheses that others are more accurate than the self for more observable, more desirable, and more automatic behaviors against people's lay theories.

**Study 3.2**

The second study directly tested the accuracy of self and other knowledge for a broad range of everyday behaviors. This study differed from previous research in two important ways. First, instead of predicting behavior in the laboratory as most previous research has done, I attempted to predict natural, real-world behavior. Second, unlike the few studies that have attempted to predict real-world behavior, I measured actual behavior instead of using self-reports of behavior as the criterion  Measures of actual behavior in Study 3.2 were obtained from participants' recordings on the Electronically

65

Activated Record (EAR; Mehl et al., 2001). This is the first study to examine the accuracy of self- and other-predictions against actual real-world behavior.

The primary purpose of Study 3.2 was to test my prediction that the domains of self- and other-knowledge are overlapping but not identical (P3.1). To test this prediction I examined whether there were behaviors that only the self was accurate in predicting, behaviors that only others were accurate in predicting, and behaviors that both the self and others predicted accurately. Once again, I also tested my three more specific predictions: that others would be more accurate than the self for more observable, more desirable, and more automatic behaviors (P3.2a-P3.2c). I also explored the results for clues to other possible moderators of self- and other-knowledge.

**STUDY 3.1**

**Method**

*Participants*

Participants were 61 residents of Austin, Texas recruited through a convenience sample. Eight research assistants each collected data from four to twelve participants whom they knew personally. Fifty-seven percent of the participants were female, the mean age was 26 (SD = 9.9 years), and the ethnic breakdown was as follows: 59% White, 26% Asian, and 15% Hispanic.

*Measure*

A scale was developed specifically for this study. On this scale, each behavior was rated on two seven-point Likert-type scales. On the first rating, participants were asked to rate "how accurate people are at predicting how much they themselves perform this behavior." On the second rating, participants were asked to rate "how accurate people are at predicting how much other people they know well perform this behavior."

The actual items were taken from a list of everyday behaviors that can reliably be coded in studies of actual behavior (see Study 3.2). This scale is referred to as the ACT (Appendix 3.1).

*Procedure*

Participants were approached by one of the research assistants in my lab and asked to take a few minutes to complete the questionnaire. The responses were anonymous – I did not ask for any identifying information. However, because the questions were not at all personal, I simply asked the participants to return the questionnaire directly to the research assistant.

**Results**

Table 4 presents the average rating of self-accuracy and other-accuracy for each behavior, as well as the difference between these two means. As can easily be seen, every behavior was rated as significantly more accurately perceived by the self than by others, except for laughing, which was rated relatively high on accuracy for both perspectives. There are large differences, however, in the perceived accuracy with which both the self and others can predict different behaviors. For example, on a scale of one to seven, participants rated the accuracy of self-predictions as a 6.07 for "attending lecture" as a 6.07 compared to 4.79 for "singing." Similarly, participants rated other-accuracy as a 5.11 for "laughing" but only 3.10 for "crying."

Table 4: Lay Perceptions of Accuracy of Self- and Other-Predictions

| Behavior | Self-Accuracy | Other-Accuracy | Difference | t |
|---|---|---|---|---|
| Crying | 5.44 | 3.10 | 2.34 | 7.46 |
| Studying/reading | 5.57 | 3.31 | 2.26 | 7.43 |
| Spend time indoors | 5.49 | 3.59 | 1.90 | 8.83 |
| Commute/in transit | 4.98 | 3.13 | 1.85 | 7.53 |
| Talking one-on-one | 5.38 | 3.52 | 1.85 | 7.24 |
| Watching TV | 5.54 | 3.70 | 1.84 | 6.71 |
| Playing sports/exercising | 5.77 | 3.97 | 1.80 | 6.96 |
| Talking on the phone | 5.31 | 3.57 | 1.74 | 5.69 |
| On the computer | 5.59 | 3.92 | 1.67 | 6.59 |
| Talking to same-sex | 5.46 | 3.80 | 1.66 | 6.47 |
| Eating | 5.21 | 3.56 | 1.66 | 6.39 |
| Talking to opposite-sex | 5.57 | 3.95 | 1.62 | 6.65 |
| Attending lecture | 6.07 | 4.49 | 1.57 | 5.66 |
| Spend time outdoors | 4.92 | 3.39 | 1.52 | 7.55 |
| Listening to music | 5.41 | 3.95 | 1.46 | 6.05 |
| Arguing | 5.03 | 3.64 | 1.39 | 4.45 |
| At a coffeeshop/bar/restaurant | 5.23 | 3.85 | 1.38 | 4.89 |
| Talking in a group | 5.28 | 4.05 | 1.23 | 5.15 |
| Social activities/entertainment | 5.31 | 4.13 | 1.18 | 5.34 |
| Spending time with others | 5.33 | 4.16 | 1.16 | 4.56 |
| Singing | 4.79 | 3.70 | 1.08 | 3.60 |
| Working at a job | 5.93 | 5.02 | 0.92 | 3.23 |
| Laughing | 5.15 | 5.11 | 0.03 | 0.16 |

*Note.* All t-values are significant at p < .01 except for "Laughing" which is not significant.

I next tested whether lay perceptions of self- and other-accuracy were related to the observability, desirability, and automaticity of the behaviors. To do this, I correlated the self- and other-accuracy ratings from the participants with experts' ratings of the observability, desirability, and automaticity of each behavior. As Table 5 shows, self-accuracy was most strongly related to the desirability and automaticity of the behaviors (more desirable and less automatic behaviors were rated higher on self-accuracy) and other-accuracy was most strongly related to observability and desirability (more

observable and desirable behaviors were rated higher on other-accuracy). Observability and desirability also predicted the difference between self- and other-accuracy such that more observable and desirable behaviors receive less discrepant accuracy ratings for self and other than did less observable and desirable behaviors.

Table 5: Correlations between Lay Predictions of Self- and Other-Accuracy and Behavior Observability, Desirability, and Automaticity

|  | Self-accuracy | Other-accuracy | Difference |
|---|---|---|---|
| Observability | .17 | .63** | -.56** |
| Desirability | .45* | .67** | -.42* |
| Automaticity | -.49* | -.19 | -.11 |

*Note.* N = 23 behaviors. * p < .05, two-tailed; ** p < .01, two-tailed.

**Discussion**

The results from this study suggest that laypeople hold the belief that the self is more accurate than others in predicting the vast majority of everyday behaviors. Laypeople believe that the self is especially good at predicting observable, intentional behaviors. According to these ratings, there is little, if anything, that close others know about our behavior that we don't know ourselves. If anything, others can somewhat predict our observable and desirable behaviors (perhaps because undesirable behaviors, such as crying, are often also less observable) about as well as we ourselves can, but they have no unique insight. The purpose of the next study was to empirically test the accuracy of self- and other-predictions of behavior. Is it true that the self knows everything that can be known about behavior?

**STUDY 3.2**

**Method**

The data from Study 3.2 presented here are from the same dataset as Study 2.1 in Chapter II. This study was a subset of a larger study. I provide a description of the

participants, as well as the measures and procedures relevant to the analyses presented in this chapter.

*Participants*

Participants were 80 undergraduate students recruited mainly from Introductory Psychology courses and flyers in the Psychology department. The sample was 54% female, 65% White, 21% Asian, 11% Latino, and 3% of another ethnicity. The vast majority of the participants (73%) were 18 years old, although their ages ranged from 18 to 24 (M = 18.7, SD = 1.4). Participants received $50 in compensation for complete participation in the large-scale study (all participants completed the entire study).

*Informants*

Each participant was asked to nominate three people who knew them well to provide ratings of their personality. Participants were asked to nominate one friend, one parent, and one romantic partner if possible. If participants could not provide any one of the three kinds of informants, they were told to nominate someone who knows them well. Participants were told that the informants' ratings would be kept completely confidential, and that they themselves would never see their informants' ratings.

Two months after the end of the study, 76% of informants had completed the ratings, resulting in a total of 182 informant ratings, of which 87 were from friends, 55 from parents, and 21 from romantic partners. The remaining informant reports were from siblings (12), ex-romantic partners (2), and one cousin, one grandmother, one great-grandmother, one ex-friend, and one informant who did not indicate his or her relationship to the target.

*Behavior*

Behavior was measured using the Electronically Activated Recorder (EAR; Mehl et al., 2001). The EAR is a digital audio recorder attached to a microphone that records the sounds of people's daily lives. The recorder can comfortably be carried in a pocket or purse, and the microphone, attached by a wire, can be worn on a shirt collar. The digital recorder was programmed to be on for 30 seconds every 12.5 minutes, producing roughly five intervals per hour. Participants could not know when the recorder was on or off. For further details on the development and testing of the device, see Mehl et al..

*Measures*

Participants and informants completed a battery of measures, which included the ACT questionnaire (Appendix 3.1). This questionnaire was designed specifically for this study, to obtain ratings of the behaviors that were subsequently measured with the EAR. Thus, the behaviors rated on the ACT are meant to represent a broad range of everyday behaviors that can be coded reliably from auditory recordings. Informants also rated the targets on ACT. All ratings were made on a 7-point Likert-type scale.

*Procedure*

All self-report measures were obtained on the first day of the three-week study. After consenting to participate, participants completed the informant-nomination questionnaire, followed by a battery of self-report questionnaires, including the ACT. The self-report questionnaires were administered online, on one of the laboratory computers. Participants were seated in a private room with the door closed, and were told to complete the questionnaires online, taking as many breaks as they desired.

As recommended by Vazire (in press-a), informant-ratings were collected via the Internet. Informants were contacted by e-mail and asked to complete an online

questionnaire about how they see the target participants' personality. Informants received a link and unique identifying number in the email. Informants who did not complete the ratings were sent reminder emails after two weeks, four weeks, and six weeks. Participants were compensated at the end of the three weeks, regardless of whether the informants had completed their ratings. Informants were not compensated for their cooperation.

Finally, participants were asked to wear the EAR for four days, starting on a Friday evening. Participants were told that the EAR would record only 4% of their daily lives, and would only capture brief snippets of conversation rather than complete conversations. Immediately after the four-day EAR recording period, participants returned to the lab and completed a questionnaire regarding their experience with the EAR. Upon completion of the study, all participants were given the opportunity to listen to their EAR recordings and erase any parts they wished.

### *Data Preparation*

Participants' behavior was coded from the ambient sounds captured on the EAR recordings. Research assistants listened to the complete recordings and, for each 30-second interval, coded the behaviors and social environments according to the Social Environment Coding of Sound Inventory (SECSI) categories developed by Mehl and Pennebaker (2003). The SECSI is a coding system that comprises the person's current location (e.g., indoors, outdoors, commuting), activity (e.g., listening to music, on the computer, eating), and interaction (e.g., alone, on the phone, talking to others).

In addition to acoustic cues such as the noise of a running engine (in transit), the sound of wind blowing (outdoors), typing noises (computer), or the voice of a lecturer (lecture), judges used context information from previous and consecutive intervals to increase their accuracy. For example, if a person, after being on campus (Interval 1) and

72

riding on a bus (Interval 2), enters an apartment (Interval 3), it is inferred that the student has returned home. The accuracy of the coding is then further enhanced by the information from the subsequent recording periods, in which the person might have switched on the TV or gone to the refrigerator to get something to eat.

**Results**

***P3.1:Domains of self- and other-knowledge are overlapping but not identical.***

Table 6 presents the accuracy of self- and other-predictions of the behaviors in descending order of accuracy for each perspective. Recall that I predicted that there would be some behaviors for which informant- predictions would be more accurate than self-predictions, and some behaviors for which self-predictions would be more accurate than informant-predictions. This prediction was supported. Self-predictions were significantly more accurate than informant-predictions for arguing ($t = 2.65$, $p < .05$) and marginally significantly more accurate than informant-predictions for watching TV ($t = 1.80$, $p < .10$). Informant-predictions were significantly more accurate than self-predictions for talking one-on-one ($t = 2.06$, $p < .05$) and attending lecture ($t = 2.63$, $p < .05$), and marginally significantly more accurate than self-predictions for spending time with others (vs. alone; $t = 1.83$, $p < .10$). Figure 3 shows the relative accuracy of the two perspectives on all 23 behaviors in the framework of the model presented above.

Overall, each perspective was accurate (at the $p < .05$ level) for 13 of the 23 behaviors. Accuracy was higher for self-predictions than informant predictions for 11 of the 23 behaviors, and higher for informant-predictions than self-predictions for 12 of the 23 behaviors. The average accuracy of across the 23 behaviors was .23 for self-predictions and .24 for informant-predictions.

Table 6: Accuracy of Self- and Informant-Predictions of Behavior

| Accuracy of Self-Predictions | r N = 79 | Accuracy of Informant-Predictions | r N = 77 |
|---|---|---|---|
| 1. Watching TV | .55** | 1. Talking on the phone | .40** |
| 2. Listening to music | .40** | 2. Watching TV | .39** |
| 3. Talking on the phone | .37** | 3. Spending time with others | .36** |
| 4. Singing | .34** | 4. Working at a job | .35** |
| 5. Talking to same-sex | .34** | 5. Listening to music | .34** |
| 6. Talking to opposite-sex | .31** | 6. Attending lecture | .33** |
| 7. On the computer | .29* | 7. Talking to opposite-sex | .32** |
| 8. Arguing | .28* | 8. On the computer | .31** |
| 9. At a coffeeshop/restaurant | .27* | 9. Social activities | .30** |
| 10. Commuting/in transit | .27* | 10. Singing | .29* |
| 11. Working at a job | .25* | 11. Talking to same-sex | .25* |
| 12. Talking in a group | .25* | 12. Laughing | .25* |
| 13. Laughing | .23* | 13. Talking one-on-one | .25* |
| 14. Social activities | .18 | 14. Talking in a group | .20$^\dagger$ |
| 15. Crying | .18 | 15. Commuting/in transit | .16 |
| 16. Spending time indoors | .16 | 16. Crying | .16 |
| 17. Spending time with others | .14 | 17. Spending time indoors | .16 |
| 18. Studying/reading | .13 | 18. At a coffeeshop/restaurant | .15 |
| 19. Spending time outdoors | .11 | 19. Playing sports/exercising | .14 |
| 20. Attending lecture | .07 | 20. Eating | .11 |
| 21. Playing sports/exercising | .06 | 21. Studying/reading | .05 |
| 22. Eating | .00 | 22. Spending time outdoors | .05 |
| 23. Talking one-on-one | -.06 | 23. Arguing | -.05 |
| Mean | .23 | Mean | .24 |

*Note.* Correlations are between predictions and actual EAR-coded behavior. ** $p < .01$, * $p < .05$, $^\dagger$ $p < .10$, all two-tailed. Behaviors in bold are predicted significantly better by one perspective than the other, $p < .10$, two-tailed.

Note that there were also many commonalities across the two perspectives. For example, both were very accurate in predicting watching TV, talking on the phone, and listening to music, and both were inaccurate in predicting eating, spending time outdoors, and playing sports/exercising. This may reflect actual overlap in domains of self- and other-knowledge, or it may reflect statistical artifacts. For example, the reliability of the behavior codings may have limited the accuracy of the predictions. To test this, I

74

correlated the accuracy of self- and informant-predictions with the reliabilities of the behavior codings. In addition, some behaviors may not have varied much across individuals, thus making it difficult to achieve accuracy. To test this, I correlated the accuracy of self- and informant-predictions across behaviors with the amount of inter-individual variance in each behavior. However, even after taking into account these statistical artifacts, there was still some similarity in the pattern of accuracy for the two perspectives. Thus, consistent with my prediction, the two perspectives had both overlapping and unique domains of knowledge.

To provide a further test of how much unique information each perspective provided, I ran multiple regressions with both informant- and self-predictions predicting actual behavior. The results, presented in Table 7, reflect the unique contribution of each perspective above and beyond the overlap between the two perspectives. For example, both self- and informant-predictions of listening to music contributed uniquely to predicting actual amount of listening to music. Both perspectives also contributed uniquely to accurate predictions of singing. Self-predictions provided unique accurate information for predictions of talking to people of the same-sex, arguing, watching TV, and commuting. Informant-predictions provided unique accurate information for predictions of spending time with others, talking one-on-one, talking on the phone, attending lecture, working at a job, and participating in social activities. These analyses provide even stronger support for my prediction that the domains self- and other-knowledge are overlapping but not identical. There are some behaviors that informants were better able to predict, and some that the self was better able to predict, and both perspectives contributed unique information that the other alone did not.

Table 7: Standardized Betas for Regression of Actual Behavior on Self- and Informant-Predictions of Behavior.

| | Self | Informants | Multiple R |
|---|---|---|---|
| Watching TV | .50** | .11 | .57 |
| Talking on the phone | .23* | .30* | .46 |
| Listening to music | .32** | .23* | .46 |
| Spending time with others | .11 | .37** | .41 |
| Singing | .29* | .22* | .40 |
| Talking to same-sex | .29* | .18 | .38 |
| Attending lecture | -.17 | .45** | .38 |
| Talking to opposite-sex | .21$^†$ | .23$^†$ | .37 |
| On the computer | .18 | .22$^†$ | .35 |
| Working at a job | .00 | .35** | .35 |
| Social activities/entertainment | .13 | .27* | .33 |
| Arguing | .33** | -.17 | .31 |
| At a coffeeshop/bar/restaurant | .27* | .02 | .28 |
| Talking in a group | .20$^†$ | .13 | .28 |
| Laughing | .14 | .20$^†$ | .28 |
| Commute/in transit | .25* | .07 | .28 |
| Talking one-on-one | -.08 | .26* | .27 |
| Crying | .16 | .08 | .21 |
| Spend time indoors | .10 | .13 | .19 |
| Playing sports/exercising | .10 | .07 | .16 |
| Studying/reading | .15 | -.02 | .14 |
| Eating | -.05 | .12 | .11 |
| Spend time outdoors | .10 | .01 | .11 |
| Means | .16 | .17 | .31 |

*Note.* N = 77. For each behavior, self- and informant-predictions were entered together into a regression predicting actual EAR-coded behavior. ** p < .01, * p < .05, $^†$ p < .10, all two-tailed.

Figure 3: Model of Self- and Other-Knowledge with Results from Study 3.2
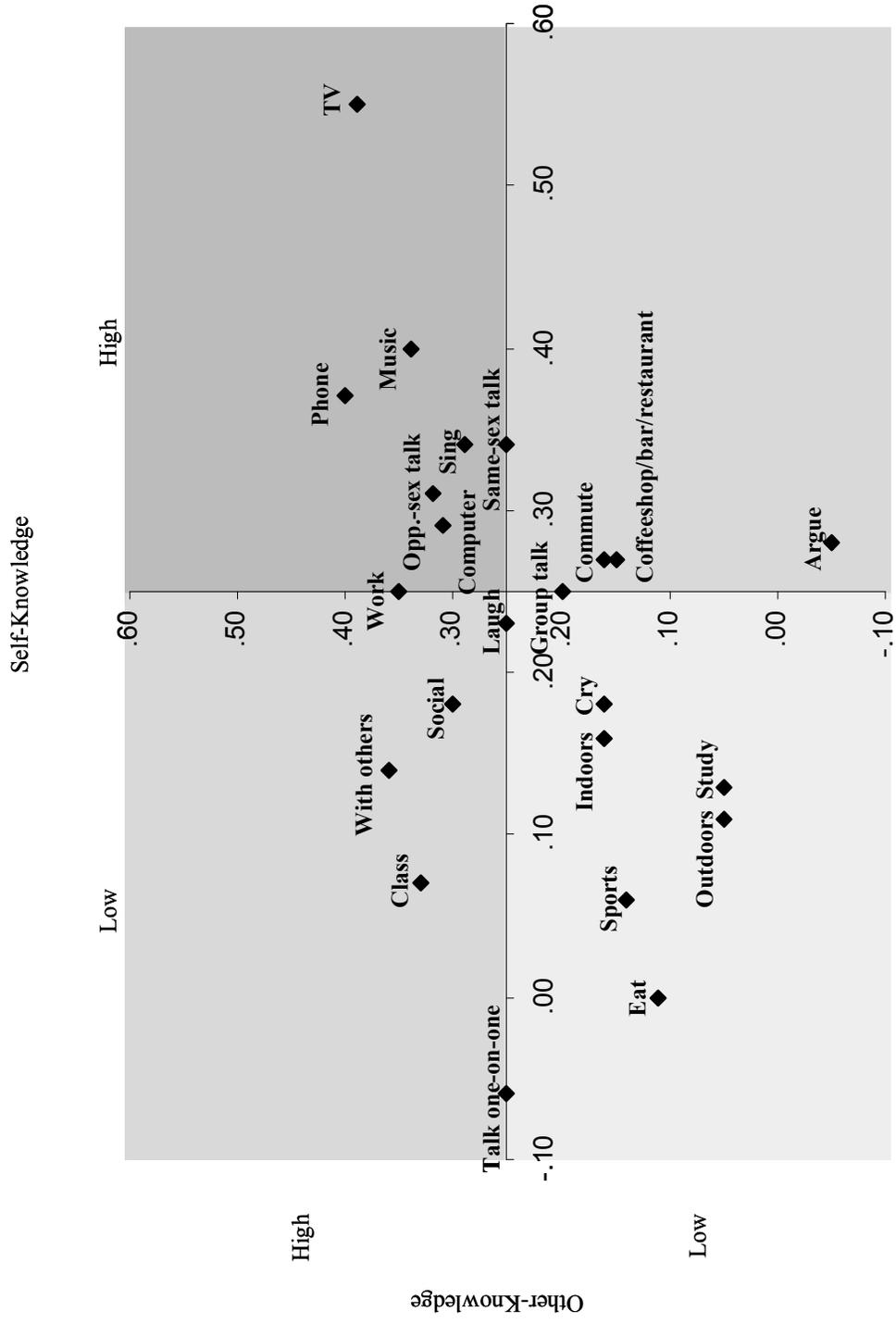
***P3.2a-3.2c: Observability, desirability, and automaticity of behaviors will be related to self-
and informant-accuracy.***

I next tested whether self- and informant-accuracy were related to the observability, desirability, and automaticity of the behaviors. To do this, I correlated the self- and other-accuracy correlations (converted to a linear variable using Fisher's r-to-z formula) with experts' ratings of the observability, desirability, and automaticity of each behavior. As Table 8 shows, self-accuracy was not related to any of these dimensions of the behaviors. Informant-accuracy was moderately related to observability (r = .30, p = .17), desirability (r = .40, p = .06), and automaticity (r = -.27, p = .20). However, due to the small number of behaviors examined in this study (N = 23), none of these correlations reached statistical significance.

Table 8: Correlations between Self- and Other-Accuracy and Behavior Observability,
Desirability, and Automaticity

|  | Self-accuracy | Informant-accuracy |
| --- | --- | --- |
| Observability | .01 | .30 |
| Desirability | -.10 | .40[†] |
| Automaticity | .09 | -.27 |

*Note.* N = 23 behaviors. [†] p < .10, two-tailed.

**Discussion**

This study tackled the issue of behavioral prediction head on. One way to examine who knows what about a person is to identify who can predict what kinds of behaviors. To test this, I examined the accuracy of self- and other-predictions of 23 everyday behaviors. The purpose of this study was to begin to map the landscape of self- and other-knowledge.

This study improved on existing studies in several ways. First, I obtained actual measures of behavior rather than self-reports or traces of behavior. This allows for a better test of accuracy because it provides a more reliable and more valid criterion measure. Second, the behaviors I examined were actual everyday behaviors aggregated over four days. This allowed me to capture individuals' natural behavioral trends. Although there are some drawbacks to measuring natural behavior rather than lab-based behavior, one clear advantage is that the results are sure to be

ecologically valid. That is, based on these findings we can be certain that self and informant ratings can be used to predict how a person will behave in his or her natural environment. Another advantage of real-world behavior is that it captures people in the environments they select and create for themselves. Laboratory-based tests, in contrast, may place people in situations that they would never encounter in their real lives.

The unique ecological design of this study allowed me to make several discoveries. First, the fact that the behavioral predictions were at all accurate strongly contradicts the situationist position that personality does not predict behavior and that behavior is not consistent across situations (e.g., Mischel, 1968). The fact that self- and informant-ratings of how the targets typically behave predicted the targets' behavior after the ratings were collected demonstrates that people do have noticeable and stable behavioral tendencies. In order to be able to predict a person's future behavior, the rater draws on global personality impressions or past behavior. In either case, the accuracy of these predictions confirms the validity of these mental constructs and the consistency of behavior across time and situations.

Another important finding is that there were sizeable differences in the magnitude of the accuracy correlations across the 23 behaviors. In some cases, this probably reflects limitations of the EAR for reliably detecting certain behaviors (e.g., playing sports). However, for other behaviors it probably reflects genuine differences in the actual predictability of the behaviors. For example, it is likely that watching TV is easier to predict than arguing or crying.

More interesting, however, are the differences in accuracy between the two perspectives. Consistent with my prediction, there were some substantial differences between self-accuracy and informant-accuracy for some traits. For example, the self was able to predict arguing better than the informants, but the reverse was true for attending lecture. Perhaps the most important conclusion that can be drawn from this finding is that, contrary to lay perceptions (Study 3.1) and many researchers' beliefs, the self is not always more accurate than others are. The simple fact that there were any behaviors that others were able to predict better than the self should cause some personality researchers to rethink their exclusive reliance on self-reports. This finding also

has far-reaching implications in the applied domain. Those trying to predict future outcomes or behaviors should consider the possibility that the outsider's perspective may be better than the self's.

The accuracy of the informants' predictions is all the more impressive when one considers that the informants were not necessarily people who live in the same town or interact with the target regularly. Many of the informants were family members and hometown friends (all participants were college students enrolled in summer school for the duration of the study). For these people, the behavioral predictions must have been based on impressions of the target in a context other than the context in which behavior was measured. For example, parents' ratings were probably based on how the participants behave at home, or on the parents' global impressions of their children's behavior.

The self had another important advantage over the informants in this study. In addition to being familiar with the context in which their behavior was recorded, people had the potential to control their own behavior. Informants who interact with the target can also manipulate or influence the target's behavior, but ultimately the self has the most control over its own behavior. This gives the self-predictions a leg-up in the accuracy domain. People can bring about the accuracy of their self-views by acting according to how they see themselves. They can also have a more subtle influence on their behavior by selecting environments or evoking behavior in others that confirm their self-views.

In short, the findings from this study show that adding the perspective of the person from the outside (i.e., informant reports) often increases the validity of behavioral predictions. Even when both perspectives were roughly equally accurate, using both as predictors substantially increased the validity of the predictions over either one alone.

Finally, this study made a first stab at trying to explain the patterns of self- and other-knowledge found in these data. By correlation the accuracy correlations with independently-rated characteristics of the behaviors, I was able to make some sense of the pattern of results. For example, I found that informants' predictions were more accurate for behaviors that are more

desirable and, to a lesser extent, for behaviors that are more observable. Although no significant patterns emerged from these analyses for self-knowledge, it is likely that this is due to the relatively small number of behaviors that were examined in this study.

Indeed, the study was less than ideal in terms of the range, number, and psychological significance of the behaviors I examined. Ideally I would have liked to measure a broad range of behaviors that reflect people's psychological traits and states. However, I was limited by the methods I have available and the desire not to abuse the good will of my participants. Thus, I could only capture thin slices of behavior over a short period of time, and I only coded behaviors that were detectable on short audio segments. Had I been able to capture a broader range of behaviors, I may have been able to provide a more detailed account of self- and other-knowledge, and may have uncovered other characteristics of behaviors that help explain these patterns. As technology improves and we become able to capture more behavior in more varied contexts, the landscape of self- and other-knowledge can be filled in with greater depth and precision.

Another limitation of this study is that because the behaviors were measured in people's natural environments, this limits the breadth of behaviors I could capture. For example, people probably select and manipulate their environments to allow them to express certain personality, and avoid expressing others (Buss, 1987). In a laboratory context, I would be able to place people in a wide range of situations that would bring out a greater variety of traits. Such a controlled setting would also allow me to examine specific behaviors that are theoretically linked to traits. In the present study it would have been difficult to generalize from the 23 behaviors to personality traits because the behaviors were not selected on the basis of their links to traits. In a laboratory setting, however, I would be able to test for specific behaviors that are known to be associated with certain traits.

Nevertheless, the results from this study provide a solid foundation from which to continue exploring these questions. Having established that self- and other-perceptions can accurately predict behavioral trends, the next step is to provide a stronger test of accuracy. As

Epstein (1983) has argued, behavioral trends are easier to predict than single behavioral acts. Furthermore, it is important to show that personality ratings, not just ratings of behavior, can predict behavioral outcome measures. Thus, the next chapter shifts to higher standard of accuracy by examining whether self- and other-perceptions of personality can predict single instances of behavior in a laboratory setting.

## Chapter 4: Accuracy of Self, Friend, and Stranger Perceptions of Personality

What does it mean to know a person? Is simply being able to predict their everyday behavior enough? Intuitively, it seems that knowing someone entails more than merely being able to describe their behavior, it entails extrapolating from that behavior to describe a person's general tendencies or predispositions. Indeed, we often describe our friends and acquaintances with personality trait terms (e.g., polite, nice, outgoing) rather than listing specific behaviors. Thus, the ultimate test of how well people know themselves and each other is to test the accuracy of their personality descriptions. I touched on this issue in Chapter 2. However, the study presented here provides a more thorough test of the accuracy of self, friend, and stranger perceptions of personality across a broader range of psychologically meaningful personality traits.

The purpose of the current study was to provide the most robust test of the accuracy of self and other judgments possible. Whereas the studies presented thus far only addressed the question of the accuracy of personality judgments indirectly or incompletely, the present study aimed to tackle the issue head on. In designing the study, I took special care to obtain the best possible measures of self- and other-views. I also chose to focus on personality traits that have been shown to have important influences in life outcomes such as physical and mental well-being, relationship satisfaction and stability, and academic and occupational success. Finally, I selected criterion measures for these traits based on validated procedures found in the literature, using multi-method indicators whenever possible.

Although there is little research directly comparing the accuracy of self and others' ratings in predicting psychologically meaningful outcomes, there is some provocative evidence that others sometimes know things that the self is unaware of. For example, in the relationship literature, it has been shown that a couple's friends are more accurate at predicting whether the couple will stay together or break up than are the members of the couple themselves (Agnew, Loving, & Drigotas, 2001). In the health literature, it has been shown that spouse's predictions of

recovery after heart failure are more accurate than self-predictions (Rohrbaugh et al., 2004). Finally, in the arena of work and academics, it has been shown that friends' ratings of conscientiousness are a better predictor of subsequent college GPA than are self-ratings of conscientiousness (Wagerman, Greve, Wright, & Funder, 2004). These and other similar findings suggest that looking beyond the self for predictors of important behavioral and life outcomes could lead to important and useful discoveries.

## DESIGN OF THE STUDY

To obtain the best possible self- and friend-ratings, I used a round-robin design in which groups of five friends rated each other's personalities and their own. The round-robin design has several important advantages over the traditional one-target-multiple-informants design used in the previous studies I have presented. The most important advantage of the round-robin design is that because each person serves as both a judge of themselves and a judge of others, we can obtain a precise estimate of how people see themselves differently from how they see others. For example, if Aaron rates himself very positively, but we don't know how he would rate others, we don't know whether Aaron has a uniquely positive view of himself or whether he sees everyone positively, including himself. However, if we ask Aaron to rate himself as well as four of his friends, Ben, Christine, Danny, and Erik, we can determine how Aaron sees himself controlling for how he sees people in general. Similarly, by having all five friends rate themselves and each other, we can determine how Aaron's friends see him differently than they tend to see everyone else. In other words, the round-robin design allows us to separate how a specific person is perceived from how people in general are perceived. The round-robin design also has the added benefit of ensuring that each participant has an equal, and high, number of informants (in this case, four friends).

To examine how the friends' perspective might differ from acquaintances', I also collected round-robin ratings for each participant from acquaintances who had just met each other during the experiment. One limitation of the previous studies (and most studies involving informant reports) is that the informants are typically self-selected. That is, people choose who

they get to know well, so the most knowledgeable informants are almost always people that the target likes and who like the target. This provides a biased picture of the third-person perspective. Ideally I would like to get highly knowledgeable informants who were randomly selected to get to know the targets well. There is potential for this kind of research in close-knit work settings (e.g., faculty members in the same department, work groups) or assigned living areas (e.g., dormitories, army barracks). However, with a typical undergraduate student sample recruited for a laboratory study, these options are unfeasible. Thus, the next best option was to introduce participants to strangers and have them get to know each other for a short period of time during the laboratory experiment. Although this level of acquaintance is far less than the friend groups', it nevertheless gives us a different third-person perspective. Presumably, the barely-acquainted strangers should be less knowledgeable about certain traits (those that are not immediately observable), but more neutral observers than the friends. Thus, collecting both friend and stranger ratings enabled me to identify the unique strengths and weaknesses of both close and distant others. In addition, including stranger ratings in my study broadened the ecological validity of the study because in many real-world settings we are called on to form impressions of strangers.

**Trait selection**

In selecting the traits to assess, I drew on literatures related to three major areas of life: love, work, and health. In the domain of love, the traits that have been found to have the most impact on mate selection, relationship satisfaction, and relationship stability include: attractiveness, neuroticism (emotional stability, positive and negative emotionality), agreeableness (kindness, likeability), extraversion (status, dominance, assertiveness), and intelligence (Botwin, Buss, & Shackelford, 1997; Buss, 1989; Buss et al., 1990; Buss, Shackelford, Kirkpatrick, & Larsen, 2001; Buunk et al., 2002; Kenrick et al., 1993; Lopes, Salovey, & Straus, 2003; Robins, Caspi, & Moffitt, 2000; Slatcher & Vazire, 2006; Slatcher, Vazire, & Pennebaker, 2006; Sprecher & Regan, 2002).

In the domain of work, the traits that have been found to have the most impact on occupational success, academic success, and job satisfaction include: conscientiousness, neuroticism (emotional stability, adjustment), extraversion (leadership, assertiveness, sociability), agreeableness, and openness to experience (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Hogan & Hogan, 2001; Hogan & Holland, 2003; Ones & Viswesvaran, 2001).

Finally, in the domain of health, the traits that have been found to have the most impact on mental health, physical health, and subjective well-being include: conscientiousness, extraversion (positive emotionality, cheerfulness), neuroticism (depression, self-esteem), intelligence, and narcissism (Danner, Snowdon, & Friesen, 2001; Friedman et al., 1995a, 1995b; Marshall et al., 1994; Martin et al., 1995; Tucker, Friedman, Wingard, & Schwartz, 1996).

Based on these three literatures, I selected the traits that are important in multiple domains and for which it is feasible to obtain a criterion measure. Thus, the traits examined in this study are: extraversion (global extraversion, dominance, and leadership ability), agreeableness (sympathy/liking of others and likeability), neuroticism (anxiety and depression), openness to experience (intelligence and creativity), narcissism (arrogance and need for power), and attractiveness.

**Criterion measures**

To evaluate the accuracy of the self-, friend-, and stranger-ratings of each of these traits, it is necessary to obtain criterion measures for each trait. The ideal criterion measure would have the following characteristics: be independent from self-views (identity) or others' views (reputation), be measured reliably (i.e., be aggregated over multiple behaviors/trials), and be conceptually very closely related to the trait for which it serves as a criterion. Of course, research is limited by many practical concerns, and not all of these standards can easily be met for each criterion measure. In the present study, however, every effort was made to maximize the objectivity, reliability, and conceptual relevance of each criterion measure. Next I describe how the criterion measures were obtained for each trait.

*Extraversion*

Like all of the Big Five personality traits, Extraversion has many facets. As mentioned above, I chose to focus on three aspects of extraversion: global extraversion (e.g., talkativeness), dominance, and leadership ability. The criterion measures for all of these traits were based on behavior during a Leaderless Group Discussion (LGD; Bass, 1954). This task was selected on the basis that it is used to assess leadership, assertiveness, and social skills in occupational settings (e.g., in employment interviews). The group members were unacquainted with each other at the beginning of the task. Details of the task are provided in the Method section below. Global extraversion was measured by having trained coders watch the videos and provide behavior counts of several extraversion-related behaviors, namely, the amount of time spent talking, number of speech acts, number of interruptions, and whether or not the participant was the one to write down the group's final decision. The coders also provided ratings of extraversion, talkativeness, and loudness which also served as criterion measures for global extraversion. The criterion measures for dominance included the behavior counts from the LGD and the coders' ratings of assertiveness, dominance, and loudness. Leadership ability was measured by asking the trained coders to rank each member of each group with respect to their leadership and contribution to the group task. Together, these traits touch on both the sociability and dominance aspects of extraversion, although of course many more facets of extraversion lie outside of these traits.

*Agreeableness*

One aspect of agreeableness is the tendency to see the best in others rather than be critical, to like others. Another aspect of agreeableness is being likeable. This is a result of agreeable people's warmth, sympathy, and conciliatory nature. I examined both of these aspects of agreeableness. The criterion measure for liking was the extent to which the person reported liking his or her friends, and the extent to which the person rated the strangers in his or her group positively. The criterion measure for likeability was the extent to which the person was liked by a group of strangers after a brief interaction, and the extent to which the person was rated as

likeable by these strangers. This criterion can be used to examine the accuracy of self- and friend-ratings of likeability, but because the criterion is based on strangers' ratings, it cannot be used to assess the accuracy of strangers' ratings. Again, these two criterion measures only tap into a fraction of the characteristics that compose agreeableness, but the tendency to like others and be liked by others is very central to the trait of agreeableness.

*Neuroticism*

Neuroticism has many distinct facets that would be difficult to assess with a single indicator. For example, depression is very distinct from anger, and it would be difficult to find one indicator that would be sensitive to both. Thus, I chose to focus on two facets of neuroticism: anxiety and depression. I chose these facets because they are strongly related to the three domains of life that were the basis for my trait selections (love, work, and health), and because there are well-validated tests of anxiety and depression. The criterion measure for anxiety was based on a variation on the Trier Social Stress Test (TSST; Kirschbaum, Pirke, & Hellhammer, 1993), which involves asking people to give an impromptu speech on a difficult topic, and telling them that their performance will be evaluated by a panel of experts. A detailed description of the procedure used is provided in the Method section below. From the speeches, trained coders provided behavior counts of anxiety-related behaviors, namely, nervous mouth movements and nervous hand movements (Egloff & Schmukle, 2002). The coders also provided ratings of each person's public speaking ability, anxiousness/nervousness, social awkwardness, expressions of insecurity and negative aspects of self, physical manifestations of tension, how relaxed/comfortable the person seemed, how fluent and quickly they spoke, how much they seemed to enjoy the task, and their overall neuroticism. These behavior codings and ratings served as criterion measures for anxiety. In addition, I also collected ratings of nervousness from the experimenter who administered the speech, self-reports of positive and negative affect immediately after the speech, and self-ratings of how anxious each person was during the speech. These also served as criterion measures for anxiety.

The criterion measure for depression was a modified version of the Beck Depression Inventory (Beck, Rial, & Rickels, 1974). This measure is not ideal because it relies on self-ratings. However, because depression is by nature an internal trait, one that is defined by subjective experience, self-ratings are arguably the best criterion available, at least among adults who are able to reflect and report on their subjective feelings.

### Openness to experience

Openness is arguably the broadest of the five personality factors of the Big Five. Thus, selecting a criterion once again necessitated narrowing the factor down to a few specific traits. I chose to focus on intelligence and creativity. This was based in part on the importance of these two traits in the literature reviewed above, and on the fact that there are well-validated laboratory-based tests of both intelligence and creativity. Intelligence was measured using the Wonderlic personnel test (Wonderlic, 1983). The Wonderlic has both verbal and non-verbal components, thus criterion measures were obtained for both of these subtypes of intelligence. The criterion measure for creativity was a timed task asking for creative solutions to a simple problem (described in detail in the Method section). The answers were scored for creativity by six research assistants.

### Narcissism

I examined two aspects of narcissism: arrogance and need for power. Both have been shown to be associated with narcissism in previous research (Vazire & Funder, in press). The short form of the Narcissistic Personality Inventory (NPI: Raskin & Terry, 1988; short form: Ames, Rose, & Anderson, in press) was used as a criterion measure for both traits. Like the criterion measure for depression, this measure is not ideal because it relies on a self-rating. However, the NPI is designed as a test rather than a face-valid self-report. Thus, it can be considered more objective than a direct self-perception. In addition to the NPI, observer's ratings of arrogance and need for power from the LGD and the TSST were used as criterion measures for each respective trait.

*Attractiveness*

The criterion measure for attractiveness was based on zero-acquaintance ratings of overall physical attractiveness, face attractiveness, and body attractiveness. The judges who provided the criterion measures had never seen the target participants and rated their attractiveness on the basis of a still photograph.

**Research questions and predictions**

The primary research question in this study is what differences exist between the self's, friends', and strangers' perspectives on personality with respect to accuracy. Who knows what about a person's personality? My prediction with respect to accuracy was that there would be differences among the perspectives. That is, I expected that the self would be more accurate for some traits, friends for other traits, and strangers for still other traits. Based on the findings presented in Chapters 2 and 3, I was able to make slightly more specific predictions. First, I predicted that the self would be especially accurate at predicting the more private traits (e.g., anxiety) and deliberate traits (e.g., friendliness), and that friends and strangers would be especially accurate at predicting the more observable traits (e.g., extraversion, attractiveness). I also predicted that the friends and strangers would be slightly more accurate overall because of aggregation over multiple observers (see Chapter 1 for a discussion of the benefits of aggregation). Because of the exploratory nature of this study, more specific predictions were not made for each of the 17 traits.

A related question is to what extent do the self, friends, and strangers have unique knowledge about each trait? That is, if more than one perspective is accurate on a trait, do the accurate perspectives know the same thing or different things about that trait? I will test this question by examining the degree to which using multiple perspectives as predictors increases the accuracy of the trait ratings.

A final question I will examine with respect to accuracy is whether patterns of accuracy in all three perspectives can be explained by three characteristics of the traits: observability, desirability, and automaticity. Here I do have specific predictions. I predict that the self will be

more accurate for less observable and less automatic traits, whereas friends and strangers will be more accurate for more observable and more automatic traits. These predictions are based on previous research discussed in Chapter 1 (John & Robins, 1993; Wilson, 2002).

In addition to examining accuracy, I will also use this study as an opportunity to examine several other research questions. First, I will examine the degree of consensus among the friends and among the strangers. Based on previous research (Funder & Colvin, 1988; Kenny, 1994), I expect that consensus will be higher among strangers than among friends. Although I expect consensus to be above chance levels, I expect it to be far from perfect. I also expect consensus to be higher for more observable traits.

Second, I will examine the degree of self-other agreement for friends and for strangers. Based on previous research (Funder & Colvin, 1988; Kenny, 1994; Watson, Hubbard, & Wiese, 2000) I expect that friends will agree more strongly with self-ratings than will strangers. I also expect that self-other agreement will be higher for more observable traits.

**METHOD**

**Participants**

Two hundred participants (40 groups of five) were scheduled to participate, but due to the threat of an impending hurricane during the last weekend of the experiment, seven groups (35 participants) either cancelled or did not show up for their scheduled session. Thus, a total of thirty-three groups of five well-acquainted friends participated in this study (N = 165). All participants were enrolled as undergraduate students at The University of Texas at Austin. Seventy-six participants (46%) were enrolled in Introductory Psychology. One hundred participants (61%) were female. Of those who indicated their ethnicity, 70 (43%) reported being Asian, Asian-American, Indian, or Pacific Islander, 65 (40%) participants reported being White or Caucasian, 18 (11%) reported being Hispanic or Latino/a, 9 (5%) reported being Black or African-American, and 2 (1%) indicated "other." Participants received $10 in compensation for the participation, were entered into a lottery with a 12% chance of winning $100, and received

partial fulfillment of course requirements if they were enrolled in Introductory Psychology. Participants were also told they would be given feedback about their personality when the study and analyses were completed.

**Procedure**

Participants were run in one of eight experimental sessions that lasted three hours and were run on Saturday and Sunday afternoons in September of 2005. The sessions each included between three and five groups of five well-acquainted friends (15 to 25 participants) and were divided into three phases, each roughly an hour long.

*Phase One*

Upon arriving, each group was taken by an experimenter to a group room. Each group room had a table with five chairs around it, and two video cameras arranged so that each participant was clearly captured on at least one camera. The cameras were not used in phase one. After explaining the study to participants and asking them to remove their hats (to avoid obscuring their faces in the video tapings), the experimenters assigned each participant a four-digit ID number, which participants wore on a sticker. Participants then provided informed consent to participate in the study, indicated whether they allowed us to use their records in presentations, and indicated whether I could contact them in the future to obtain follow-up measures.

Next, participants were given the round-robin rating forms as well as manila folders to be used to hide their answers from their group members. The round-robin rating forms can be found in Appendix 4.1. Participants were given approximately 15 minutes to rate all five members of their group (including themselves). While participants completed the round-robin ratings, the experimenter also rated each participant on the same scale.

During the last five minutes of the ratings, the experimenter initiated the first saliva-collection procedure. Each participant was given a cup of water to rinse their mouth, and asked to chew a piece of gum for five minutes. At the end of the five minutes, when the round-robin

ratings were completed, all participants provided a saliva sample by drooling into a vial through a straw. The vials, labeled with the participants' ID numbers, were collected immediately by an experimenter and stored in a freezer.

### Phase Two

At the end of phase one, all participants were reunited in the lobby area and were asked to divide themselves according to the color on their ID sticker. The stickers were color-coded such that the phase two stranger groups included one participant from each of the phase one friend groups. The experimenters took their new stranger groups back to one of the group rooms. Experimenters verified that none of the participants knew each other (none ever did).

Pizza and soda were delivered to each group room, and the experimenter explained that the group would have eight minutes to eat and talk, and were told that they should use this time to try to get to know each other as well as possible. They were told that they would be asked to rate each other's personalities after the get-to-know-you interaction. The cameras were turned on and the experimenter left the room. After eight minutes, the experimenter returned and turned off the cameras. Participants then completed round-robin ratings identical to the ones completed in their friend groups in phase one, with the exception that they did not rate themselves again in phase two.

### Phase Three

After phase two, the participants were once again reunited in the lobby area and asked to divide themselves based on the shape on their ID stickers. This process resulted in new stranger groups that were once again comprised of one member from each of the original friend groups and resulted in groups of previously unacquainted participants. Each group was taken to one of five stations and every fifteen minutes, the groups were rotated such that after 75 minutes, all five groups had been to every station.

One station consisted of the Leaderless Group Discussion (LGD; Bass, 1954.). Because the groups in phase three varied in size from three to five members (depending on the number of

friend groups present for that session), the procedure for the LGD varied slightly. In all cases, the participants were told that they would be participating in a group discussion in which each person is a representative from a different branch of the same company and they are meeting to allocate the end of the year merit bonuses for their employees. Participants were told that their two goals were to obtain as much of a bonus for their own employee as possible, and to help the committee decide the best allocation of the available funds. In addition, it was emphasized that this task is commonly used to assess leadership ability in job interviews, and that they should treat this activity as seriously as they would a job interview. The complete instructions for the five-member groups can be found in Appendix 4.2. After giving the instructions, the experimenter turned on the video camera (which captured all participants). Participants then had one minute to read over the description of the employees in the running for the merit bonus, and then were given ten minutes to discuss and come to a final decision. In the middle of the table, equidistant from the participants, was a form on which participants were told to write down their final decision. The experimenter left the room for the duration of the discussion. When one minute was left, the experimenter warned the group that they must come to a decision in the next minute, and write it down on the form given to them. If the group came to a decision before the ten minutes were up, the experimenter did not stop the interaction. After ten minutes, the experimenter returned, collected the form, and turned off the video camera. Participants were then asked to complete a written ranking of how much each group member contributed to the task, including themselves. These rankings were done confidentially behind manila folders so that the group members could not see each other's rankings.

Another station consisted of a modified version of the Trier Social Stress Test (TSST; Kirschbaum, Pirke, & Hellhammer, 1993). At this station, participants completed the Positive and Negative Affect Scales (PANAS; Watson, Clark, & Tellegen, 1988) before being told what task they would be doing. After completing the PANAS, participants were taken individually into a cramped room with a video camera and a stern experimenter. They were told that they would have to give a speech which would be evaluated by a panel of experts. It was emphasized

that public speaking ability is an important predictor of many life outcomes, including both occupational and interpersonal success. Participants were told they would be given ten seconds to prepare after being told the subject of the speech, and that they would then have to talk continuously for two minutes. They were then told that the topic of their speech would be "what I like and don't like about my body." They were given ten seconds to prepare and then asked to begin. The speech was videotaped. The experimenters were trained not to smile, laugh, or otherwise put the participants at ease. If the participants asked any questions after the speech had started, the experimenters simply said "please continue." If the participants stopped before the two minutes were over, the experimenter asked the participant to continue. After two minutes, the experimenter abruptly cut off the participant and turned the video camera off. Participants were then taken to another room, given a stopwatch to indicate the time that had elapsed since the speech, and were asked to complete the PANAS again. When their stopwatch indicated that 15 minutes had passed since the end of their speech, another saliva sample was collected. The experimenter who administered the speech rated each participant's nervousness immediately after the participant left the speech room.

Another station consisted of the Picture Story Exercise (PSE). The PSE consists of asking participants to write an imaginative story about each of six pictures: ship captain, couple by river, trapeze artists, women in laboratory (from Smith, 1992), boxer (from McClelland & Steele, 1972), and nightclub scene (from McClelland, 1975). Pictures were presented using standard instructions and procedures described in Smith (1992). The full instructions can be found in Appendix 4.3. Participants were given four minutes for each picture, and completed four pictures at this station and two more at another station.

Another station consisted of the 12-minute Wonderlic IQ test (Wonderlic, 1983), a 60-second creativity test asking participants to list as many uses as they could think of for a brick (Friedman & Förster, 2002) and the Over-Claiming Questionnaire (OCQ; Paulhus, Harms, Bruce, & Lysy, 2003).

Finally, another station consisted of the last two pictures of the PSE and a photo session. In the photo session, participants were photographed in a natural stance and a standardized stance (arms shoulder-width apart, hands at sides, and a neutral facial expression). Participants were also videotaped jumping. After completing all stations, the groups were debriefed and paid.

**Coding**

*LGD*

A team of four coders coded the LGD discussions. Each group discussion was coded by two independent coders who watched the video three times for each participant. After the first viewing of a target participant, the coder completed the same 40-item personality rating that was completed by the participants in their round-robin groups, and an additional 19 trait and behavior items, all rated on a 15-point Likert-type scale. During the second viewing of a participant the coder recorded whether the participant was the first to speak, how many times the participant interrupted someone, and whether the participant was the one to write down the group's final decision. During the third viewing the coder recorded the number of speeches made by the participant and the length of each speech. A speech act was defined as any time the participant took the floor. This excluded short exclamations such as "ok" or "yeah." The two coders' data were aggregated into a single composite. All reliabilities (Cronbach's α) were .62 or higher.

*TSST*

A team of four coders coded the TSST speeches. Each speech was coded by two independent coders who watched the video four times for each participant. After the first viewing, the coder completed the same 59-item rating that the LGD coders completed. During the second viewing, the coder counted the number of nervous mouth movements, defined as lip biting, lip twitching, licking of the mouth, and biting of the lips. During the third viewing, the coder counted the number of eye blinks. After the fourth viewing, the coder rated the amount of anxiety signaled through hand movements and position on a 15-point Likert-type scale. All three of these behaviors have been used as indicators of anxiety in previous research (Egloff &

Schmukle, 2002). The two coders' data were aggregated into a single composite. The cross-coder reliability (Cronbach's α) of the aggregate of the anxiety behaviors coded from the TSS was .68.

*Creativity*

All answers to the brick test (Friedman & Förster, 2002) were scored for creativity by six independent judges. The judges were members of the research team. Each response was given a creativity score based on the aggregate of the judges' responses (α = .64). Each participant was then given a creativity score based on the mean creativity rating of the uses they listed.

*Attractiveness*

A team of three coders viewed still photographs of each participant and rated each participant's overall attractiveness, face attractiveness, and body attractiveness. Two coders were male, one was female. The three coders' data were aggregated into a single composite (α = .72 for overall attractiveness, .65 for face attractiveness, and .73 for body attractiveness).

**RESULTS**

**Accuracy**

Accuracy was measured by correlating the self, friend, and stranger ratings with the criterion measures for each trait. Because the ratings were based on only a few items and the behavioral outcome measures were based on a single situation, both the predictor and the criterion were less reliable than I would have ideally liked. In order to improve the reliability of the criterion variable, I computed the accuracy of the ratings by comparing them to the aggregate of all the relevant behavioral criterion measures. This provides a more reliable and more accurate estimate of the actual effect size. However, I also report the correlations between the ratings and each individual behavior. Despite the fact that these correlations are often small in magnitude, I will refer to ratings as "accurate" if the effect size is statistically significant. This usually corresponds to an effect size of at least .20. Although this effect size is small, it is not corrected for the unreliability of the measures, and thus is probably an underestimate of the actual accuracy. Furthermore, an effect of .20 or .30 would be very practically significant if the ratings

were used to predict repeated or aggregated behavior, as personality judgments are usually used in real life. Thus, my discussion of the results assumes that these effect sizes are practically significant.

In addition, a great deal of power is necessary to detect significant differences between two dependent correlations. Thus, the alpha level used in the statistical tests for the differences in accuracy between perspectives is .10 rather than the standard .05. Furthermore, I occasionally discuss differences that are not significant if they are large or if there is a consistent pattern across traits.

### Extraversion

#### Global extraversion

Table 9 presents the accuracy correlations for ratings of extraversion. The ratings are a composite of the two extraversion items on the Ten Item Personality Inventory (Gosling, Rentfrow, & Swann, 2003): "Is extraverted, enthusiastic" and "Is reserved, quiet" (reverse-scored). The reliability (α) of the two-item aggregate is .69, .69, and .64 for self, friends, and strangers respectively. The criterion measures are all derived from codings of the LGD. As Table 9 shows, self, friend, and stranger perceptions of extraversion all correlated with behavior during the LGD. Extraversion was especially correlated with the number of times participants interrupted someone else during the LGD, and with how extraverted, talkative, and loud they were perceived as being by the LGD coders. The aggregate accuracy correlations for the self, friend, and stranger ratings were all about .25. Considering that the ratings were based on only two items and the criterion was usually a single item, these correlations are quite impressive. This is especially true for the strangers' perceptions, which were based on only an eight-minute interaction.

Table 9: Correlations between Self, Friend, and Stranger Ratings of Extraversion and Criterion Measures

|  | Self | Friends | Strangers |
|---|---|---|---|
| Time talking | .20 | .14 | .26 |
| Number of speeches | .18 | .16 | .15 |
| Interruptions | .23 | .22 | .25 |
| Final decision | .13 | .09 | .19 |
| Rank in group | .10 | .10 | .19 |
| Extraverted | .26 | .24 | .21 |
| Talkative | .26 | .28 | .27 |
| Loud | .24 | .20 | .23 |
| Mean | .20 | .18 | .22 |
| Aggregate | .25 | .22 | .26 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Criterion measures are derived from codings of the LGD discussions by two trained coders. Correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

**Dominance**

Table 10 presents the accuracy correlations for ratings of dominance. The ratings are based on the single item "Tends to dominate group discussions". As Table 10 shows, self, friend, and stranger perceptions of dominance all correlated with behavior during the LGD. Dominance was especially correlated with the amount of talking during the LGD, with the LGD coders' ranking of their contribution to the LGD task, and with how assertive, dominant, and loud they were perceived as being by the LGD coders. The aggregate accuracy correlations for all three perspectives were about .30. Again, although these correlations are small to medium in absolute magnitude, it is impressive to find such a relationship between two-item personality ratings and behavior in a laboratory task.

Table 10: Correlations between Self, Friend, and Stranger Ratings of Dominance and Criterion
Measures

|  | Self | Friends | Strangers |
|---|---|---|---|
| Time talking | .24 | .20 | .27 |
| Number of speeches | .16 | .21 | .25 |
| Interruptions | .20 | .20 | .29 |
| Final decision | .20 | .20 | .29 |
| Rank in group | .27 | .26 | .22 |
| Assertive | .28 | .31 | .29 |
| Dominates group discussions | .27 | .29 | .22 |
| Loud | .32 | .29 | .28 |
| Mean | .24 | .25 | .26 |
| Aggregate | .31 | .31 | .34 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Criterion measures are derived from codings of the LGD discussions by two trained coders. Correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

**Leadership**

Table 11 presents the accuracy correlations for ratings of leadership. The ratings are based on the single item "Is a good leader." As Table 11 shows, self, friend, and stranger ratings of leadership correlated with LGD coders' ratings of both contribution to the LGD task and leadership ability (r's around .20). Across the three extraversion traits, a clear pattern emerges: all three perspectives are accurate at predicting extraversion-related behaviors. There is no clear advantage of the self's perspective or of acquaintance (friends); strangers were just as good at rating extraversion as anyone else.

Table 11: Correlations between Self, Friend, and Stranger Ratings of Leadership and Criterion
      Measures

|               | Self | Friends | Strangers |
|---------------|------|---------|-----------|
| Rank in group | .14  | .23     | .21       |
| Good leader   | .15  | .21     | .16       |
| Mean          | .15  | .22     | .19       |
| Aggregate     | .16  | .24     | .21       |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Criterion measures are derived from codings of the LGD discussions by two trained coders. Correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.


## *Agreeableness*

### **Sympathy/liking**

Table 12 presents the accuracy correlations for ratings of liking. The ratings are based on the single item "tends to like others." However, instead of using raw scores on this item, the self ratings were based on the self effects and friends' and strangers' ratings were based on target effects for each participant. The self and target effects are derived from Social Relations Model analyses. Self effects represent how a person rated him or herself controlling for how they tend to rate people in general and how others tend to rate them (i.e., how they uniquely see themselves). Target effects represent how a person is seen by others (either friends or strangers) controlling for how people in general are seen in their group (i.e., how they uniquely are seen).

As Table 12 shows, ratings of how much the participant tends to like others correlated with both how much the participant reported liking their friends and the strangers in their group, and how positively they rated their friends and the strangers. Interestingly, the friends' ratings predicted the outcomes (liking and positivity) in the friend groups (mean r = .32) but not in the stranger groups (mean r = .10). Self ratings predicted liking and positivity in friend groups (mean r = .18) more than in stranger groups (mean r = .12), and stranger ratings did not significantly predict any of the outcomes. As the aggregate accuracy correlations show, only the self and

friends were accurate at predicting how much the target liked others. Friends were slightly more accurate than the self, but this difference was not significant.

Table 12: Correlations between Self, Friend, and Stranger Ratings of Liking Others and Criterion Measures

|  | Self | Friends | Strangers |
|---|---|---|---|
| Liking of friends | .17 a | .27 a | -.04 b |
| Positivity of ratings of friends | .18 a | .37 b | -.06 c |
| Liking of strangers | .17 | .02 | .09 |
| Positivity of ratings of strangers | .06 | .17 | .12 |
| Mean | .15 | .21 | .03 |
| Aggregate | .20 | .29 | .04 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Self ratings are the self effects saved from SRM analyses of the round-robin data. Friends' ratings are the target effects saved from SRM analyses of the friend round-robin data. Strangers' ratings are the target effects saved from SRM analyses of the stranger round-robin data. The criterion measures are perceiver effects saved from SRM analyses of the friend and stranger round-robin data. Correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

**Likeability**

Table 13 presents the accuracy correlations for ratings of likeability. The ratings are based on an aggregate of the items "Is likeable" and "Tends to be liked by others." The reliability (α) of the two-item aggregate is .76 and .89 for self and friends respectively. Accuracy could not be computed for strangers because the strangers' ratings served as the criterion. The first criterion measure is the strangers' actual liking of each participant (the single item: "How much do you like this person?"). The second criterion is the strangers' ratings on the same two-item likeability aggregate ("Is likeable" and "Tends to be liked by others" α = .83). As Table 13 shows, self and friend ratings of likeability both predicted how much participants were liked and found likeable by the strangers in their groups (aggregate r's are around .20 to .30). Friends' ratings were slightly more predictive, but the difference in accuracy between self and friends was not significant. However, this pattern was also found for liking, suggesting that friends may be slightly more accurate than the self at judging agreeableness-related traits.

Table 13: Correlations between Self, Friend, and Stranger Ratings of Likeability and Criterion Measures

|                         | Self | Friends |
|-------------------------|------|---------|
| Liked by strangers      | .18  | .22     |
| Likeability - strangers | .16  | .26     |
| Mean                    | .17  | .24     |
| Aggregate               | .19  | .27     |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Accuracy of likeability ratings could only be computed for self and friends because the strangers' ratings of liking and likeability served as the criterion measures. Correlations in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

### *Neuroticism*

#### **Anxiety**

Table 14 presents the accuracy correlations for ratings of anxiety. Two separate ratings were used as predictors of anxiety. First, the neuroticism ratings are a composite of the two neuroticism items on the Ten Item Personality Inventory (Gosling, Rentfrow, & Swann, 2003): "Is anxious, easily upset" and "Is calm, emotionally stable" (reverse-scored). The reliability (α) of the two-item aggregate is .62, .56, and .00 for self, friends, and strangers respectively. However, because this two-item aggregate is not very reliable (especially among the strangers' ratings) and is much broader than the criterion measure (anxiety during a public speaking exercise), I also examined the accuracy of ratings on the single item "Is good at public speaking." The criterion measures in the top half of Table 14 are derived from codings of the TSST speeches. The accuracy correlations for the aggregate of these 13 variables (α = .91) is presented in the bottom row of the top half of Table. The criterion measures in the bottom half of Table 14 are the ratings of nervousness by the experimenter who administered the speech and self-ratings of affect immediately after the speech. The bottom row of the bottom half of Table 6 presents the mean accuracy correlation for these items.

As Table 14 shows, ratings of neuroticism did not correlate with any of the TSST-based criterion measures of anxiety (mean r's are .05, .05, and -.02 for self, friends, and strangers,

respectively), or with RA- and self-rated anxiety (mean r's are .15, .02, and -.06 for self, friends, and strangers). The only striking exceptions are the accuracy of self-rated neuroticism in predicting coders' ratings of how negative the participants were about themselves during the speech (r = .22) and predicting self-rated negative affect immediately after the speech (r = .33).

Ratings of public speaking ability, however, were much more useful for predicting anxiety. All three perspectives accurately predicted the TSST-coded criterion measures (r's with aggregate are .43, .35, and .24 for self, friends, and strangers respectively). A test of significance of difference between dependent correlations shows that self-ratings were significantly more accurate at predicting TSST anxiety than were stranger ratings (t = 1.97, p < .05). Ratings of public speaking were especially accurate at predicting the TSST coders' ratings of public speaking ability, awkwardness of interpersonal style (reversed), fluency and speed of talking, and how much the participant seemed to enjoy the task. Self-ratings of public speaking were also more accurate than friend or stranger ratings in predicting RA- and self-rated anxiety during the speech (mean r's are .34, .17, and .01 for self, friends, and strangers respectively; $t_{self, friends}$ = 2.40, p < .05, $t_{self, strangers}$ = 3.50, p < .01).

Table 14: Correlations between Self, Friend, and Stranger Ratings of Anxiety and Criterion Measures

| | Neuroticism | | | Public Speaking | | |
|---|---|---|---|---|---|---|
| | Self | Friends | Strangers | Self | Friends | Strangers |
| Nervous mouth movements | -.12 | -.12 | -.11 | .24 a | .14 a | .02 b |
| Nervous hand movements | -.11 | -.01 | -.12 | .31 a | .08 b | .06 b |
| Public speaking | -.01 | .02 | -.01 | .43 a | .34 a | .17 b |
| Anxious/nervous | .01 | .01 | -.09 | .36 a | .19 b | .17 b |
| Awkward | .11 a | -.03 ab | -.14 b | .44 a | .35 ab | .24 b |
| Insecurity | .16 a | .18 a | -.10 b | .27 | .24 | .20 |
| Negative about self | .22 | .10 | .06 | .23 ab | .32 a | .14 b |
| Physical tension | .04 | .02 | -.03 | .29 a | .17 ab | .08 b |
| Relaxed/comfortable ® | -.01 | -.03 | -.06 | .34 | .22 | .18 |
| Fluent ® | -.02 | -.04 | -.01 | .44 a | .37 a | .20 b |
| Speaks quickly ® | -.10 | -.07 | .02 | .37 | .34 | .27 |
| Enjoys interaction ® | .13 a | .07 a | -.11 b | .35 | .32 | .24 |
| Neurotic | -.01 a | .16 b | .16 ab | .04 | .01 | .05 |
| Aggregate | .05 | .05 | -.02 | .43 a | .35 ab | .24 b |
| RA rated nervousness | .06 | .02 | -.08 | .32 a | .23 a | .05 b |
| PA post speech ® | .06 | .03 | -.07 | .29 a | .06 b | .18 ab |
| NA post speech | .33 a | .08 b | -.08 a | .29 a | .12 b | -.12 c |
| Anxious post-speech | .13 a | -.06 b | -.01 a | .51 a | .26 b | -.06 c |
| Mean a | .15 | .02 | -.06 | .36 a | .17 b | .01 c |

Note. N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. a Mean is the mean of the correlations in the last four rows of the table. Correlations within each trait in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

105

**Depression**

Table 15 presents the accuracy correlations for ratings of depression. The ratings of depression are based on the aggregate of the items "Is anxious, easily upset," "Is calm, emotionally stable" (reversed), "Is happy, satisfied with life" (reversed), "Is lonely," "Has high self-esteem" (reversed), and "Is depressed." The reliability ($\alpha$) of the six-item aggregate was .78, .80, and .60 for self, friends, and strangers respectively. The criterion measure for depression was the participants' scores on the Beck Depression Inventory short-form (BDI; Beck, Rial, & Rickels, 1974), with one item deleted (item 7 on suicidal ideation) resulting in a 12-item questionnaire ($\alpha = .68$). As Table 15 shows, self-ratings of depression were significantly more accurate than friend- and stranger-ratings (t = 3.08, p < .01), and friend ratings were significantly more accurate than stranger ratings (t = 3.38, p < .01). Interestingly, the differences among the three perspectives were especially accentuated among female targets. This is due mostly to the fact that female participants' self-ratings correlated more strongly with their scores on the BDI than did males'.

Table 15: Correlations between Self, Friend, and Stranger Ratings of Depression and the Criterion Measure

|  | Self | Friends | Strangers |
|---|---|---|---|
| BDI | .52 a | .28 b | -.06 c |
| males only | .35 a | .26 a | -.12 b |
| females only | .63 a | .31 b | -.01 c |

*Note.* N = 165. Male N = 65. Female N = 100. All r's ≥ .20 are significant at p < .01, two-tailed. BDI = Beck Depression Inventory (short form) without item 7 (suicidal ideation). Correlations within each trait in the same row with different subscripts are significantly different from each other at an $\alpha$ level of .10, two-tailed.

*Openness to Experience*

**Intelligence**

Table 16 presents the accuracy correlations for ratings of intelligence. The ratings are single-item scores on three different items: "Is intelligent," "Has strong verbal skills,"

and "Has strong math skills." The three items were not aggregated because separate criterion measures were used for each item. The criterion measure for "Is intelligent" was participants' overall score on the Wonderlic Personnel Test (Wonderlic, 1983). The criterion measures for the other two items were the participants' scores on the verbal and non-verbal subsections of the Wonderlic. As Table 16 shows, self and friends' ratings were significantly more accurate than strangers' ($t_{self-stranger}$ = 1.86, p < .10, $t_{friend-stranger}$ = 2.54, p < .05). The magnitude of the correlations found here are typical of those reported in the literature (Borkenau & Liebler, 1993; Reynolds & Gifford, 2001).

Table 16: Correlations between Self, Friend, and Stranger Ratings of Intelligence and Criterion Measures

|           | Self        | Friends     | Strangers    |
|-----------|-------------|-------------|--------------|
| IQ        | .07 a       | .23 b       | .00 a        |
| Verbal    | .22 a       | .22 a       | -.05 b       |
| Math      | .18 ab      | .27 a       | .02 b        |
| Mean      | .16 a       | .24 a       | -.01 b       |
| Aggregate | .19 a       | .28 a       | -.02 b       |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. IQ correlations are correlations between ratings of "intelligence" and overall IQ score on the Wonderlic Personnel Test. Verbal correlations are correlations between ratings of "has strong verbal skills" and scores on the verbal section of the Wonderlic. Math correlations are correlations between ratings of "has strong math skills" and scores on the non-verbal section of the Wonderlic. Correlations within each trait in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

**Creativity**

Table 17 presents the accuracy correlations for ratings of creativity. The ratings are based on the single item "Thinks and associates ideas in unusual ways, has unconventional thought processes." The criterion measure was the participants' scores on the brick creativity test described above (Friedman & Förster, 2002). As Table 17 shows, all three perspectives' ratings of creativity correlated somewhat with participants' scores

on the creativity test. This is impressive considering that the creativity test was a single 60-second test, and that creativity is more subject to idiosyncratic interpretation than many other traits (e.g., facial attractiveness).

Table 17: Correlations between Self, Friend, and Stranger Ratings of Creativity and the Criterion Measure

|  | Self | Friends | Strangers |
|---|---|---|---|
| Brick Test Score | .21 | .27 | .20 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Brick test score = participants' scores on the creativity test (Friedman & Förster, 2002). Correlations within each trait in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

### Narcissism

#### Arrogance

Table 18 presents the accuracy correlations for ratings of arrogance. The ratings are based on an aggregate of the items "Is arrogant" and "Exaggerates his/her abilities." The reliability (α) for the two-item aggregate was .67, .85, and .83 for self, friends, and strangers respectively. The criterion measures were: arrogance as rated by the LGD coders, arrogance as rated by the TSST speech coders, and scores on the NPI. As Table 18 shows, all three perspectives were accurate at predicting arrogance. If anything, the self was even more accurate than friends or strangers (this difference was significant for self and stranger ratings of arrogance predicting NPI scores; t = 2.26, p < .05). The accuracy of self-ratings is impressive considering that arrogance is thought of as a trait on which people are not very self-aware (i.e., arrogant people are thought to be oblivious to the fact that they are arrogant, hence the need for tests like the NPI). The accuracy of strangers' ratings is also impressive considering that they had very little information on the targets.

Table 18: Correlations between Self, Friend, and Stranger Ratings of Narcissism and
Criterion Measures

|  | Self | Friends | Strangers |
|---|---|---|---|
| Arrogance in LGD | .04 | .09 | .16 |
| Arrogance in TSST | .21 | .17 | .14 |
| NPI | .41 a | .29 ab | .19 b |
| Mean | .23 | .19 | .17 |
| Aggregate | .33 | .27 | .24 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. LGD = Leaderless group discussion. TSST = Trier Social Stress Test. NPI = Narcissistic Personality Inventory (short-form). Correlations within each trait in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

**Need for Power**

Table 19 presents the accuracy correlations for ratings of need for power. Ratings were based on the single item "Is power-oriented, values power in self and others." The criterion measures were: need for power as rated by the LGD coders, need for power as rated by the TSST speech coders, and scores on the NPI.

Table 19: Correlations between Self, Friend, and Stranger Ratings of Need for Power and
Criterion Measures

|  | Self | Friends | Strangers |
|---|---|---|---|
| nPower LGD | .06 a | .27 b | .12 ab |
| nPower TSS | .20 | .23 | .20 |
| NPI | .37 a | .33 a | -.01 b |
| Mean | .21 ab | .28 a | .10 b |
| Aggregate | .32 a | .42 a | .14 b |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. nPower = Need for power. LGD = Leaderless group discussion. TSST = Trier Social Stress Test. NPI = Narcissistic Personality Inventory (short-form). Correlations within each trait in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

As Table 19 shows, both self and friends were accurate at predicting need for power. Strangers' ratings of need for power only correlated with power-orientation

during the LGD task. Tests of significance of differences between dependent correlations shows that stranger ratings were significantly less accurate than both self (t = 3.43, p < .01) and friend (t = 2.97, p < .01) ratings.

### *Attractiveness*

**Attractiveness**

Table 20 presents the accuracy correlations for ratings of attractiveness. The ratings are single-item scores on three different items: "Is physically attractive," "Has an attractive face," and "Has an attractive body." The three items were not aggregated because separate criterion measures were used for each item. The criterion measure for "Is physically attractive" was unacquainted observers' ratings of participants' attractiveness based on a photograph. The criterion measures for the other two items were the observers' ratings of facial and body attractiveness respectively. As Table 20 shows, all three of the three perspectives were quite accurate at predicting attractiveness. However, self-ratings of face attractiveness were significantly less accurate than friend and stranger ratings ($t_{self-friend}$ = 2.02, p < .05; $t_{self-stranger}$ = 2.51, p < .05).

Table 20: Correlations between Self, Friend, and Stranger Ratings of Attractiveness and Criterion Measures

|  | Self | Friends | Strangers |
|---|---|---|---|
| Attractiveness | .37 | .48 | .44 |
| Face | .25 a | .41 b | .46 b |
| Body | .46 | .51 | .47 |
| Mean | .36 | .47 | .46 |
| Aggregate | .41 | .50 | .48 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Attractiveness correlations are correlations between ratings of "physical attractiveness" and overall attractiveness ratings from three judges who viewed still photographs of the targets. Face and body correlations are the accuracy of those specific judgments. Correlations within each trait in the same row with different subscripts are significantly different from each other at an α level of .10, two-tailed.

Having examined the accuracy of each perspective alone, I next examined how much validity the perspectives add to each other. That is, do the three perspectives' knowledge overlap completely, or is there some unique information that only one perspective has over the others? To test this, I conducted a regression for each of the 17 personality traits, regressing the criterion measure onto the three perspectives' trait ratings. Table 21 presents the beta weights for each perspective and the multiple R for the regression when all three perspectives' ratings are entered simultaneously. The raw accuracy correlations of the aggregates from tables 7 to 19 are also presented for comparison.

As Table 21 shows, there are instances of unique information across the three perspectives. For example, ratings of body attractiveness were independently predictive for all three perspectives, even when entered simultaneously into a regression. However, there were also instances of one perspective subsuming the other two, such as for depression, where the self was uniquely valid and wiped out any utility of the other two perspectives. It is also important to note that the multiple R was typically quite a bit stronger than any of the individual accuracy correlations. This indicates that for most traits, using multiple perspectives leads to greater accuracy.

Table 21: Validity of Trait Ratings: Correlations and Regression Coefficients

| | r | | | B | | | Multiple |
|---|---|---|---|---|---|---|---|
| | Self | Friends | Strangers | Self | Friends | Strangers | R |
| **Extraversion** | | | | | | | |
| Global ext. | .25 | .22 | .26 | .23 | .13 | .16 | .41 |
| Dominance | .31 | .31 | .34 | .19 | .17 | .25 | .44 |
| Leadership | .16 | .24 | .21 | .09 | .20 | .18 | .32 |
| Aggregate | .35 | .30 | .30 | .23 | .13 | .16 | .41 |
| **Agreeableness** | | | | | | | |
| Liking | .20 | .29 | .04 | .15 | .25 | -.01 | .32 |
| Likeability | .19 | .27 | -- | .12 | .23 | -- | .29 |
| Aggregate | .27 | .29 | -- | .31 | .19 | -- | .41 |
| **Neuroticism** | | | | | | | |
| Anxiety | .43 | .35 | .24 | .34 | .17 | .17 | .50 |
| Depression | .52 | .28 | -.06 | .47 | .14 | -.05 | .54 |
| Aggregate | .40 | .28 | .04 | .35 | .18 | -.03 | .43 |
| **Openness** | | | | | | | |
| Intelligence | .19 | .28 | -.02 | .11 | .25 | -.05 | .30 |
| Creativity | .21 | .27 | .20 | .15 | .21 | .17 | .35 |
| Aggregate | .20 | .30 | .09 | .14 | .25 | .08 | .33 |
| **Narcissism** | | | | | | | |
| Arrogance | .33 | .27 | .24 | .27 | .15 | .18 | .41 |
| Need for power | .32 | .42 | .14 | .23 | .35 | .09 | .48 |
| Aggregate | .37 | .34 | .20 | .30 | .22 | .13 | .45 |
| **Attractiveness** | | | | | | | |
| Overall | .37 | .48 | .44 | .19 | .28 | .24 | .56 |
| Face | .25 | .41 | .46 | .07 | .21 | .33 | .50 |
| Body | .46 | .51 | .47 | .28 | .26 | .26 | .62 |
| Aggregate | .41 | .50 | .48 | .21 | .26 | .27 | .59 |
| Mean (traits) | .30 | .33 | .23 | .22 | .21 | .14 | .43 |
| Mean (aggregates) | .33 | .34 | .23 | .26 | .21 | .12 | .44 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. All B's ≥ .18 are significant at p < .05, two-tailed; all B's ≥ .15 are significant at p < .10, two-tailed. All multiple R's significant at p < .05.

Next I examined whether the patterns of accuracy for the three perspectives can be explained in part by variations in observability, desirability, and automaticity across the traits. To test this, I correlated each perspective's accuracy correlations with the observability, desirability, and automaticity ratings of the traits (collected in Study 3.2). The results of these analyses are presented in Table 22.

Table 22: Correlations between Self-, Friend-, and Stranger-Accuracy and Trait
Observability, Desirability, and Automaticity

|                | Self  | Friends | Strangers |
|----------------|-------|---------|-----------|
| Observability  | .19   | .56     | .86       |
| Desirability   | -.56  | .13     | .39       |
| Automaticity   | -.15  | -.02    | -.16      |

*Note.* N = 17 traits. All r's ≥ .5 are significant at p < .05, two-tailed; all r's ≥ .8 are significant at p < .01, two-tailed.

As predicted, friends and strangers were more accurate on traits that are more observable. I also predicted that the self would be less accurate for more automatic behaviors. This prediction was not supported (although the trend is in the right direction). I also found that self-ratings were less accurate for more desirable traits, which I had not predicted.

**Agreement**

*Consensus*

Table 23 presents the agreement levels among friends and among strangers. Consensus was calculated in two ways. First, I calculated the average pairwise correlation among the raters (friends or strangers). Second, I used the target variance estimates derived from Social Relations Model analyses. Target variance is the proportion of variance in ratings accounted for by target effects. Target effects are tendencies for targets to be seen a certain way by perceivers. That is, social relations model analyses

separate the variance in ratings due to perceiver biases (perceiver variance) from the variance due to the actual target (target variance).

Table 23: Consensus Among Friends and Among Strangers

|  | Friends | | Strangers | |
|---|---|---|---|---|
|  | Mean pairwise r | Target effect | Mean pairwise r | Target effect |
| Extraversion | | | | |
|     Global extraversion | .40 | .39 | .45 | .49 |
|     Dominance | .33 | .34 | .37 | .41 |
|     Leadership | .24 | .24 | .21 | .22 |
| Agreeableness | | | | |
|     Liking/sympathy | .11 | .16 | .17 | .14 |
|     Likeability | .24 | .25 | .12 | .14 |
| Neuroticism | | | | |
|     Anxiety - neuroticism | .36 | .29 | .03 | .00 |
|     Anxiety - public speaking | .30 | .31 | .26 | .23 |
|     Depression | .33 | .19 | .13 | .18 |
| Openness | | | | |
|     Intelligence | .27 | .24 | .20 | .21 |
|     Verbal intelligence | .23 | .22 | .19 | .17 |
|     Math skills | .39 | .34 | .25 | .26 |
|     Creativity | .20 | .21 | .00 | .10 |
| Narcissism | | | | |
|     Arrogance | .32 | .26 | .20 | .19 |
|     Need for power | .21 | .14 | .15 | .15 |
| Attractiveness | | | | |
|     Overall | .36 | .34 | .37 | .34 |
|     Face | .33 | .32 | .37 | .35 |
|     Body | .44 | .41 | .34 | .33 |
|  MEAN | .30 | .27 | .23 | .23 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. Target effects are saved from SRM analyses of friend and stranger round-robin groups.

    As Table 23 shows, consensus was generally higher among friends than among strangers on both indicators of consensus. There were also interesting differences across

traits. For example, consensus on neuroticism was high among friends but low among strangers, whereas consensus on attractiveness and extraversion was high in both groups. This suggests that strangers only reach consensus for very observable traits, whereas friends are able to form consensual impressions even on traits that are not very observable. Overall, friends' ratings achieved significant levels of consensus ($r > .16$) for 16 of the 17 traits (all but liking/sympathy) and strangers' ratings achieved significant levels of consensus for 12 of the 17 traits (all but likeability, neuroticism, depression, creativity, and need for power).

### *Cross-perspective agreement*

Table 24 shows the agreement correlations among the three perspectives: self, friends, and strangers. Self-friend and self-stranger agreement was calculated in two ways. First, I computed the raw correlation between self-ratings and aggregated friend ratings for each trait. Second, I used the correlation between self effects and target effects in the Social Relations Model analyses. Friend-stranger agreement was calculated simply by computing the correlation between aggregated friends' ratings and aggregated strangers' ratings for each trait.

Table 24: Cross-Perspective Agreement

| | | Self-Friend | | Self-Stranger | | Friend-Stranger |
|---|---|---|---|---|---|---|
| | | r | SRM | r | SRM | r |
| **Extraversion** | | | | | | |
| | Global extraversion | .54 | .67 | .44 | .55 | .46 |
| | Dominance | .38 | .46 | .24 | .33 | .25 |
| | Leadership | .28 | .36 | .07 | .15 | .08 |
| **Agreeableness** | | | | | | |
| | Liking/sympathy | .21 | .31 | .15 | .34 | .18 |
| | Likeability | .28 | .33 | .16 | .37 | .26 |
| **Neuroticism** | | | | | | |
| | Anxiety - neuroticism | .47 | .57 | -.07 | .00 | .11 |
| | Anxiety - public speaking | .42 | .49 | .08 | .05 | .21 |
| | Depression | .34 | .38 | -.07 | .09 | .11 |
| **Openness** | | | | | | |
| | Intelligence | .20 | .29 | -.09 | -.23 | .16 |
| | Verbal intelligence | .33 | .53 | -.04 | .02 | .11 |
| | Math skills | .55 | .62 | .27 | .31 | .38 |
| | Creativity | .29 | .44 | .08 | .22 | .13 |
| **Narcissism** | | | | | | |
| | Arrogance | .30 | .33 | .09 | .12 | .22 |
| | Need for power | .24 | .36 | .05 | .17 | .06 |
| **Attractiveness** | | | | | | |
| | Overall | .40 | .39 | .29 | .32 | .53 |
| | Face | .39 | .37 | .29 | .32 | .53 |
| | Body | .41 | .42 | .29 | .31 | .52 |
| Mean | | .36 | .43 | .13 | .20 | .26 |

*Note.* N = 165. All r's ≥ .16 are significant at p < .05, two-tailed; all r's ≥ .20 are significant at p < .01, two-tailed. SRM = correlation between self-rating and SRM target effect saved from SRM analyses of friend and stranger round-robin groups.

In general, self-friend agreement was stronger than self-stranger agreement, and friend-stranger agreement fell somewhere in between the two. There were also interesting differences across traits. For example, although self-friend agreement was high for almost all traits, self-stranger agreement was only high for extraversion and attractiveness. This

could be due in part to the low levels of consensus among strangers for the other traits (see Table 23). Interestingly, friend-stranger agreement was the highest cross-perspective agreement for attractiveness. It is also interesting to note that strangers consistently agreed more with the friends than with the self for every trait. This could be due in part to the increased reliability of the friend ratings over the self-ratings (due to aggregation), but may also reflect the commonalities across various types of observer perspectives.

Recall that I predicted that both consensus and cross-perspective agreement would be higher for more observable traits. To test this prediction and examine other potential characteristics of traits that could covary with consensus and cross-perspective agreement, I correlated the consensus correlations and the cross-perspective agreement correlations with trait observability, desirability, and automaticity (collected in Study 3.2). The results from these analyses are presented in Table 25.

Table 25: Correlations between Consensus and Cross-Perspective Agreement and Trait Observability, Desirability, and Automaticity

|  | Friends Consensus | Strangers Consensus | Self-Friend Agreement | Self-Stranger Agreement | Friend-Stranger Agreement |
|---|---|---|---|---|---|
| Observability | .38 | .42$^{\dagger}$ | .21 | .65** | .74** |
| Desirability | -.12 | .10 | -.03 | .46* | .44$^{\dagger}$ |
| Automaticity | -.10 | -.47* | -.28 | -.28 | .02 |

*Note.* N = 17 traits. $^{\dagger}$ p < .10, two-tailed; * p < .05, two-tailed; ** p < .01, two-tailed.

As predicted, consensus and cross-perspective agreement was higher for more observable traits. This was especially true when strangers were involved. A few other significant correlations were found but they are hard to interpret without a theoretical basis for predicting them.

**DISCUSSION**

**Summary of findings**

The purpose of this study was to determine which perspective (self, friend, or stranger) was the most accurate at predicting a collection of traits that have been found to have psychological importance in the domains of love, work, and health. The key finding here is that no single perspective was the best across the board. For example, the self was the most accurate at predicting anxiety and depression, which are both internal and relatively unobservable, but was the least accurate at predicting attractiveness, the most observable trait examined in this study.

Another important finding with respect to accuracy is the sheer level of accuracy obtained by all three perspectives in this study. The self was able to predict outcomes related not only to extraversion and neuroticism (traits that are relatively easy for the self to judge), but also arrogance, need for power, and attractiveness – traits that are generally thought to be difficult to access by the self. Friends were able to predict every outcome with at least some accuracy. Among the 17 traits examined in this study, friends had some insight into each of them.

The accuracy of the strangers' predictions was the most remarkable – strangers were as accurate as the self and friends for some of the traits. Specifically, strangers were able to predict extraversion, dominance, leadership, arrogance, and attractiveness with about the same level of accuracy as either of the other two perspectives. This is very impressive considering that the strangers only interacted with each other for eight minutes before rating each other, and did not interact with each other at all during the criterion tests.

What could explain the accuracy of the stranger's ratings? One possibility is that people are very skilled at detecting one another's personality traits. There are several

118

reasons to suspect this. First, previous research has shown that snap judgments of personality can be very accurate, even when the judgment is based only on a photograph or other thin slices of information (Ambady & Rosenthal, 1992; Berry & Finch Wero, 1993), including seeing the target's bedroom, office, or website (Gosling, Ko, Mannarelli, & Morris, 2002; Vazire & Gosling, 2004), or even knowing their music preferences (Rentfrow & Gosling, 2003). Second, it would be evolutionarily adaptive to be able to infer a person's personality very quickly. Thus, the ability to quickly and accurately judge others' personalities has probably been selected for in our species.

Another possible explanation is that people broadcast their personality in very observable ways. There are several reasons to believe this also plays a part in the accuracy of strangers' ratings. First, there is very little evidence that there are individual differences in how well people are able to detect others' personalities. Many studies have searched for the "good judge" of personality but no consistent finding has emerged (Pickhardt, Vazire, & Gosling, 2003). Thus, it is unlikely that the accuracy is due entirely to a few good judges detecting targets' personalities. If everyone is able to pick up on one another's personality traits, these traits must be quite observable. Indeed, research has shown that extraverts are particularly easy targets of personality judgment (Colvin, 1993; Funder, 1999; Swann & Rentfrow, 2001). That is, extraverts broadcast many of their traits, not just their extraversion. Thus, the accuracy of the strangers' ratings may have been due, in part, to their accurate detection of extraverts' personalities.

The degree to which self, friend, and stranger ratings predicted single objective behavioral criterion measures was also impressive. For example, ratings of extraversion predicted the amount of time a person spoke during the leaderless group discussion and the number of times they interrupted someone during the discussion, and self-ratings of public speaking ability predicted the number of nervous mouth and hand movements

119

during a stressful speech. Although the effect sizes for the predictions of single behavioral criteria were often small, the accuracy of the aggregated criterion measures was often in the .30 to .40 range. Besides providing a substantial increase in the magnitude of the accuracy correlations, this demonstrates that the estimates obtained in this study are constrained by the reliabilities of the criterion measures (and probably the ratings as well). Had it been feasible to obtain more reliable measures, the accuracy correlations would have been even stronger.

The regression analyses showed that the three perspectives sometimes had unique knowledge about a trait. That is, even when multiple perspectives were accurate, they often accounted for different aspects of the criterion measure. This finding is important because it shows that the perspectives differ in content, not just amount of information. It also demonstrates that using multiple perspectives leads to substantially higher validities because in many cases the perspectives supplement each other rather than being redundant.

This study also provided some important insights into the patterns of self- and other-knowledge. Specifically, I found that friends and strangers know more about observable traits. Although this is not surprising, it has never directly been tested before. In addition, I also found that the self was less accurate for more desirable traits. Although researchers have examined the impact of trait evaluativeness on self-knowledge (e.g., John & Robins, 1993), the finding that self-ratings are less accurate for more desirable traits is a new one. I tested the possibility that this was an effect of evaluativeness rather than desirability, and that was not supported. This effect was indeed due to self-ratings being less accurate for desirable (but not undesirable) traits. Finally, I found a weak trend suggesting that the self may be more accurate for less automatic (i.e., more intentional) traits, and strangers may be more accurate for more automatic traits.

I also took this opportunity to examine levels of agreement within perspectives and across perspectives. Somewhat surprisingly, friends agreed more with each other than did strangers. Our results suggest that consensus increases with greater acquaintance, but only for traits that are very difficult to observe (e.g., anxiety, depression, creativity, arrogance). It is likely that strangers had very little information about these traits and thus were reduced to essentially guessing, leading to lower levels of consensus.

Agreement levels across perspectives were generally lower than agreement within perspectives. However, one notable exception is that the friends tended to agree more with the self-ratings than they did with each other. However, this could be an artifact of the increased reliability of the aggregated friends' ratings over the individual friends' ratings. To test this, I computed the average pairwise self-friend agreement. The results from these analyses confirm that this is indeed a psychometric artifact (mean pairwise self-friend $r = .25$ across 17 traits). Thus, consistent with previous research (Funder & Colvin, 1988; Funder, Kolar, & Blackman, 1995), agreement was consistently stronger within perspective than across perspectives. Also consistent with previous research (John & Robins, 1993), agreement within and across perspectives was higher for more observable traits.

**Implications for Self-Knowledge**

This study sheds light on the strengths and limitations of the self as a source of information about a person. The accuracy findings show that the self is uniquely knowledgeable about some traits, but also uniquely oblivious about others. Although only tentative conclusions can be drawn from a single study, the results suggest that the self is a uniquely accurate source of information about internal traits such as anxiety and depression. The self is also quite aware of levels of extraversion, dominance, and even

arrogance and attractiveness. This suggests that whatever biases the self may have (e.g., self-enhancement, self-verification, etc.), people are not completely blind to even highly evaluative traits such as arrogance and attractiveness. Arrogant people seem to have some idea of their arrogance, and unattractive people seem to be quite aware of their unattractiveness. Clearly, self-awareness is not debilitated by self-perception biases.

However, portraying the self as omniscient would also be mistaken. Self-ratings of traits such as likeability, intelligence, and even liking of others were only minimally accurate. The last is particularly surprising – people should have complete awareness of the degree to which they tend to like others or not. The fact that friends, and even strangers, were sometimes better at predicting outcomes than the self was shows that there are indeed obstacles to self-awareness. Further research is needed to identify the nature of these obstacles: are they due to motivated biases, informational deficits, or a mistaken impression of the self's comparison group?

**Implications for Personality Judgment**

The results from this study are very encouraging with respect to the ability of lay people to accurately read each other's personalities. Friends were able to predict every single outcome with some accuracy. This study improves on the previous studies presented in this dissertation because the friends represented a wider range of liking than in the other studies. This is because rather than nominating their own best friends as informants, participants had to come in groups of five. This led to groups having a range of liking and acquaintance among the friends. Thus, rather than everyone being rated only by the two or three people who like them most, each person was rated by four people who may or may not have been immediate friends of theirs. This results in a different picture of how people are seen by their peers. These means suggest that the friends in this study were not quite as positive in their ratings as the friends in the

previous studies. However, because positivity and accuracy are independent, it is not clear whether the reduced positivity bias of the friends' ratings has an impact on their accuracy. What is clear is that friends can, at least sometimes, report accurately on a person's level of extraversion, agreeableness, neuroticism, openness, narcissism, and attractiveness. For a broad view of what a person is like, friends might be the best source to go to.

**The role of acquaintance**

Do friends have unique insight because of their closeness with the target, or would any third-person perspective be just as accurate? The strangers' perspective provides a way to examine this question, and also to determine how much of a person's personality can be captured through a brief, superficial get-to-know you interaction, typical of those people have when they first meet, for example, at parties, on the bus, or at work. The results from this study suggest that acquaintance does matter – friends and strangers have distinctly different views of a person from each other.

The means suggest that consistent with previous findings, strangers were more harsh in their ratings than were friends or the self. It is impossible to tell whether this reflects a harshness bias or merely a lack of positivity bias. The accuracy findings suggest, however, that strangers were not more accurate than friends for any trait. That is, to the extent that friends' ratings were more positive than strangers', this did not seem to adversely affect the accuracy of their ratings. Friends were better than strangers at rating agreeableness, neuroticism, intelligence, and need for power.

What is striking about strangers' ratings, however, is that they were able to accurately predict any of the outcomes. Recall that the strangers only interacted with each other for eight minutes, and were then separated for the remainder of the study. Despite the very small amount of information they had about each other, strangers' ratings of

extraversion, dominance, leadership, arrogance, and attractiveness predicted the criterion measures for those traits. In just eight minutes, participants were able to accurately detect psychologically important traits such as assertiveness and arrogance.

**Implications for Love, Work, and Health**

The application of these findings depends on the traits that are being assessed and the assessment context. When resources are plentiful, it is clearly always best to get as many perspectives on a person as possible. However, many times this is not feasible. When can we use only strangers' ratings? When are self-ratings to be trusted? The results from this study suggest that self ratings are valid for most traits, but not all. Research has demonstrated, for example, that self-ratings of intelligence should not be used as a proxy for intelligence (Paulhus, Lysy, & Yik, 1998), and the findings from this study support that conclusion. Furthermore, people seem to not have a very good idea of how much they are liked by others or how much they tend to like others. Agreeableness has been shown to be one of the most important predictors of success in romantic relationships (and in certain types of jobs; Hogan & Holland, 2003). Thus, researchers and practitioners interested in agreeableness will need to go beyond self-reports. In contrast, people have a good idea of how assertive they are, which is also important both in relationships and in occupational contexts (Hogan & Holland, 2003). With respect to mental health, the self is the best informant – people are very good at knowing their levels of anxiety and depression. This is likely to generalize to other internal experiences such as physical health and pain.

As mentioned above, friends provide the most consistently accurate information about a person. Thus, researchers and practitioners who can afford to do so should obtain peer ratings of their subjects, ideally from peers who know the target well but do not all like the target tremendously.

124

Finally, strangers can provide very useful information on a limited range of traits. Those interested in measuring traits related to extraversion and narcissism would probably be safe if they relied only on stranger ratings. This knowledge could greatly simplify the personality assessment process in situations were those are the only relevant traits.

**Limitations**

Although extensive, this study has a number of limitations. First, the friend groups varied greatly in their composition with respect to gender, closeness among the friends, liking among the friends, and the nature of the relationships among the friends. Second, the participants were all undergraduate students at UT Austin. Thus, I cannot confidently say whether these results generalize to other age groups, geographic regions, etc. Third, due to the length and complexity of this study, participants may have lost interest and not taken the study as seriously as I would have liked.

Perhaps the most important limitation of this study is that the criterion measures were less than ideal. First, choosing objective measures that are conceptually related to the traits assessed involves some conjecture. As Funder (2001) points out, our knowledge of the links between personality and behavior is relatively thin, making it difficult to know what the best criterion measures are for each trait. Second, because this is a laboratory study, the criterion measures were somewhat artificial. For example, a better measure of anxiety would have been to observe people's reaction to a natural stressor in their lives, rather than creating a stressor in the laboratory. Third, criterion measures should ideally be aggregated over many instances and situations, rather than being based on a single event. This was not possible in the current study. However, I do plan to continue collecting data from these participants (i.e., turn this into a longitudinal study), and future waves of data collection may give me the opportunity to collect more, and

more ecologically valid, criterion measures. For example, I will eventually have long-term outcome measures such as who graduates from college, who gets a job, who gets married, etc. I would also like to continue to add new methods to the procedure. For example, I would like to collect EAR data from these participants. This would allow me to compare the participants' behavior in the lab to their behavior in their own natural environments.

# Chapter 5: General Discussion

The challenge of uncovering a person's true personality continues to beguile researchers and laypeople. Indeed, many people spend countless hours wondering about their own true personality. Do we know who we are? How do we know who a person really is?

The aim of this dissertation was to explore the depths and limits of self- and other-knowledge. What do we know about ourselves? What do others know about us that we are oblivious to? In addition to providing a basic description of the content of self- and other-knowledge, it was my hope that some of the findings presented here would illuminate the reasons behind these patterns. What kinds of things do people know about themselves? Why are there some things people seem oblivious to? The answers to these questions have important implications for both researchers and laypeople. I took several approaches to these questions in this dissertation, testing the accuracy of self- and other-perceptions with increasingly rigorous tests of accuracy.

## OVERVIEW OF THE FINDINGS

The purpose of chapter 2 was to catalogue the differences between the self and close others' views of personality, and provide a preliminary test of the accuracy of these views. Interpersonal perception research has often treated the self-other distinction as the only important distinction in the universe of possible perceivers. Some research has examined the distinction between strangers and well-acquainted others. However, this is the first study to distinguish between and systematically compare five important high-acquaintance perspectives: the self, friends, parents, romantic partners, and siblings. All of these individuals or groups play an important role in our lives, and yet we know very little about how their perceptions differ from one another.

127

The studies in chapter 2 demonstrated that there are numerous meaningful differences among the perspectives. The self is less positive than any other perspective and parents more positive. Overall self-other agreement is about the same for all perspectives, but there is some evidence that for specific traits, the perspectives differ in how much their perceptions are in line with self-perceptions. And the perspectives differ in how much they know about targets' attractiveness, self-esteem, and depression.

Although the findings presented in chapter 2 are novel and important, they are only the tip of the iceberg. It is likely that even within these informant groups, there are important differences among informants. For example, there are important differences between same-sex and opposite-sex friendships (Bleske & Buss, 2000), which are likely to influence the accuracy of personality judgment. Romantic relationships also vary greatly across dimensions such as sexual orientation and length of relationship, and these differences could also affect the accuracy of personality judgment. Future research should explore these possibilities and attempt to identify the attributes of relationships (e.g., closeness, disclosure, sexual interest, cohabitation, etc.) that may affect interpersonal perceptions.

The self-other agreement correlations presented in chapter 2 provide a very preliminary test of accuracy. Self- and other-perceptions are likely to be valid to the extent that they agree with other indicators of what a person is like. Thus, higher levels of self-other agreement suggest more potential for accuracy. I also presented a few analyses that directly tested the accuracy of the self- and other-perceptions. The results from these analyses show that there are differences across perspectives in accuracy. For example, friends seem to be best at rating attractiveness, and partners are the most accurate at rating depression. However, it was not possible to test the accuracy of ratings for most traits because no clear criterion was available. Rather than focusing on accuracy, this

chapter laid the groundwork for establishing the differences among the perspectives and testing the accuracy of self- and other-perceptions more rigorously in the next two chapters.

The purpose of chapter 3 was to provide a more direct test of self- and other-accuracy. By obtaining self- and informant-predictions of behavior and comparing them to actual behavior in participants' everyday lives, I was able to determine how accurate each perspective is at predicting 23 different behaviors.

The findings from chapter 3 support the prediction that both perspectives would have unique, accurate predictions for different behaviors. That is, the self knows about some behaviors better than the informants do, but informants in turn know more about other behaviors than the self does. Neither perspective is supreme or omniscient, each has its own domain of expertise.

The findings from this study are exceptional because of the ecological validity of the criterion measure. The results show that self- and other-predictions of behavior correspond with actual real-world behavior. Previous studies have shown that personality ratings correlate with self-reports of behavior, or with narrow behaviors collected in laboratory settings, but this study demonstrates that perceptions of a person's personality (as measured by ratings of their typical behavior) predict their actual behavior in their everyday lives. This is particularly impressive because people often select and craft their real-world environments to fit their personality, which means that certain traits and behaviors that could be elicited in the laboratory can be avoided in real life.

Given that self- and other-perceptions can predict specific behaviors, the next chapter presented a more stringent test of accuracy: can self- and other-perceptions predict single laboratory-based behavior. Chapter 3 examined the accuracy of predictions of behavioral trends. However, as Epstein (1983) has argued, predicting single behaviors

is a stronger test of accuracy than predicting behavioral trends. Furthermore, the behaviors predicted in Chapter 3 were micro-level behaviors. A more interesting question is whether self- and other-perceptions can predict psychologically meaningful behaviors and outcomes.

The purpose of chapter 4 was to examine the accuracy of self, friend, and stranger ratings of personality on psychologically important traits. I collected ratings of personality as well as laboratory-based behavioral criterion measures for a broad range of traits that are important for love, work, and health. I examined accuracy by comparing the self, friend, and stranger ratings of personality to these criterion measures.

The results show that ratings from all three perspectives are considerably accurate, at least for some traits. The findings also demonstrate the importance of aggregating across behaviors, and of matching the level of specificity of the predictions to the level of specificity of the outcome. For example, ratings of global neuroticism did not predict anxiety during the speech, but ratings of public speaking ability did.

These findings further bolster the case for personality, showing that judgments of personality do predict important objective outcomes. In fact, the three perspectives together were able to accurately predict the criterion measure for all of the traits examined, with an average multiple correlation coefficient of .43 across the 14 traits. More importantly, different traits were best predicted by different perspectives. Although strangers never outperformed the self or friends, they did manage to match their accuracy for several traits. Friends were uniquely accurate at rating agreeableness and intelligence. The self was uniquely accurate at rating neuroticism. For many of the traits, such as dominance and arrogance, all three perspectives had unique knowledge.

Taken together, the results presented in this dissertation show that perceptions of an individual vary across perceivers, and that the self has some unique advantages and

130

disadvantages. These differing perceptions reflect many things, including enhancement and diminishment biases, a focus on the pragmatic aspects personality in a given context, and the multiple realities of who a person really is. To some extent, the differences across perspectives reflect inaccuracies, but often they also may reflect actual differences in how people behave in different contexts, and the utility of perceivers being able to predict context-specific outcomes. In the following sections I discuss the specific implications of these findings for self-knowledge and interpersonal perception.

## WHAT DOES THE SELF KNOW?

The studies presented in this dissertation lead me to conclude that people know a great deal about themselves. People's perceptions of their own personalities are compatible with how they are seen by others, their predictions of their own behavior are substantially accurate, and their perceptions of their own traits correspond with their behavior. There are limits to self-knowledge, however. For example, people seem to know more about undesirable traits than about desirable ones. People seem quite aware of their levels of neuroticism, depression, and narcissism, but not of their agreeableness and intelligence. Even when it comes to their looks, people don't know nearly as much as they could about their attractiveness.

One possible explanation for this variation in self-knowledge is that the self knows about traits that refer to a person's identity, but not about traits that are social effects. For example, anxiety is a trait that resides completely within the person; it is not defined by the effect a person has on others. In contrast, agreeableness is a trait that is defined by the effect one has on others. Indeed, some of the most important and interesting aspects of personality are entirely defined by how we affect others. For example, humor, trustworthiness, charm, and sexual appeal are all social effects and have important consequences in life. Future studies should examine this characteristic of a trait

131

(whether it is defined by one's identity or one's effect on others) as a potential moderator of self-knowledge.

The finding that the self is less accurate at judging desirable traits is new and surprising. Previous research has found that evaluativeness, but not desirability, is associated with greater self-peer and peer-peer agreement, such that agreement is higher for less evaluative traits (John & Robins, 1993). This suggests that judges are better at rating neutral than evaluative traits. The findings presented here, however, suggest that the self may be good at judging undesirable traits, but not desirable ones. No strong conclusions can be drawn from these data, however, due to the small number of traits that were examined (the moderating effect of desirability may have been due to the effects of one or two extremely desirable or undesirable traits).

What are the causes of inaccuracy in self-perceptions? Although the studies presented here did not directly address this question, several insights can be drawn from the findings. First, people seem to be overly harsh in their self-ratings. There are, however, substantial individual differences in the positivity of self-ratings. This suggests that, for some people, a self-diminishment bias may be the cause of some inaccuracy. Another potential cause of inaccuracy is a simple lack of self-awareness. Contrary to Bem's (1972) self-perception theory, self-perception is not as straightforward as simply observing one's own behavior. If that were the case, self-predictions of behavior in Chapter 3 would have been at least as accurate as informant-predictions for all behaviors. This suggests that people are not as good at observing their own behavior as they could be. Future research should explore this possibility by examining the influence of feedback (e.g., showing people video tapes of their own behavior) on the accuracy of self-perception. This would allow us to disentangle inaccuracy due to lack of awareness from motivated inaccuracy (e.g., self-deception or lying).

It should be emphasized, however, that the inaccuracies in self-perception found in these studies were not debilitating. Self-perceptions still correspond strongly to other indicators of personality, and are extremely useful in predicting psychologically meaningful outcomes and behaviors. Rather than demonstrating the weaknesses of self-perception, the more important implication of the inaccuracies in self-knowledge is that they demonstrate that personality and self-perceptions are theoretically independent. Our views of who we are greatly overlap with our true personalities, but the two are conceptually and empirically distinct.

## WHAT DO OTHERS KNOW?

Like self-knowledge, our knowledge of others who are close to us is quite extensive. We agree on what their personalities are like, we can predict how they will behave with some accuracy, and our ratings of their personality traits correspond with their behavior. However, our domains of expertise of others' personalities are quite different than our domains of expertise about our own personalities. When it comes to knowing others, we know more about their observable traits than their less visible ones, and, if they are strangers, we are better at detecting desirable attributes than undesirable ones. We are quick to pick up on their levels of extraversion, creativity and attractiveness, but are less knowledgeable about their neuroticism and depression.

This pattern of findings raises several interesting questions. First, how do others come to know so much about us? One possibility is that, as discussed above, others know about those traits that are by definition social effects. For example, our peers may be more in tune with our level of charm, sense of humor, or sexual appeal than we are ourselves, because those traits are reputational in nature. It is also likely that others know a lot about us because it is very important for us to be able to predict each other's behaviors. The evolutionary pressure to understand and predict each other's behavior

133

may be responsible for the evolution of self-consciousness (which led to the ability to make predictions about our own and others' behavior; Vazire & Robins, 2004), and has probably led to a fine-tuned system for judging one another's personalities.

The value of being able to predict others' behavior, however, is limited to the context in which we are likely to encounter these specific others. As Swann (1984) has argued, we only need to know people in the circumscribed contexts in which we interact with them. Thus, we do not need to know what our friends are like with their families or lovers, nor what our siblings are like with their friends or lovers. This gives a new standard for judging the circumscribed accuracy of an informant's rating. Perhaps the informants in my studies would have been even more accurate if the outcomes they were predicting were in the same context as their relationship to the target. In other words, it is likely that friends would be especially accurate at predicting friendship-related outcomes, family members at predicting family-related outcomes, and so on. Thus, different informants will be accurate at predicting different outcomes in part because their motivations for knowing the target differ (e.g., romantic partners have more reason to know about their partner's emotional stability than do friends).

The idea of circumscribed accuracy makes sense because, pragmatically, we only need to be able to predict a person's behavior to the extent that it is relevant to our relationship with that person. However, there is another important reason to suspect that an informant's perception will be especially accurate in the circumscribed context of their relationship to the target. Specifically, informants' perceptions may differ because the target may behave differently with different informants. That is, if people behave differently with their friends than they do with their siblings, friends and siblings are likely to have different, but equally accurate, views of the same person. Thus, different informants will be accurate at predicting different outcomes in part because the actual

personality of the target may vary with different informants (e.g., people may behave more neurotically with their partners than with their friends).

The idea that people's personalities fluctuate across contexts has begun to receive increased attention from personality researchers (Fleeson, 2001; Wood & Harms, 2005; Wood & Roberts, in press). In the domain of relationships, people's context-specific personality (how they are with their partners) predict relationship satisfaction and stability better than does global personality (Slatcher & Vazire, 2006). Similarly, global well-being trickles down and affects job satisfaction and relationship satisfaction, but each of these is also predicted by context-specific affect measures (Heller & Watson, 2005; Heller, Watson, & Ilies, 2004). These and similar studies suggest that the discrepancies across perspectives found in my studies may reflect real differences in personality and behavior across contexts. This has serious implications for the interpretation of my results: disagreements across perspectives do not necessarily reflect inaccuracy, but may reflect different realities. This also suggests that informant reports can be even more useful and predictive when they are used to predict outcomes in the domain in which the informant knows the target. Just as informants as a whole provide a complementary perspective to the self's, different informants from different contexts can add incremental knowledge to our understanding of a person by adding new information about what the target is like in a different context.

IMPLICATIONS

Anyone interested in predicting the behavior of individuals, whether for research purposes, personnel selection, mate selection, or other reasons, would benefit from understanding what people do and don't know about themselves, and why. With increasing evidence showing that self-perceptions are far from perfect, we as researchers

135

are often left with a sense of helplessness – if people don't even know themselves, how can we ever hope to know them?

The same issues arise in our personal lives. Upon meeting someone, we often wonder where to turn if we want to know what that person is like or what they are likely to do in the future. Can we trust their self-reports? Should we ask mutual friends? These issues are also relevant to employers making hiring decisions, advisors selecting graduate students, lawyers selecting jury members, and many other everyday situations. Despite the prevalence of these concerns, little research has been done on the accuracy and inaccuracy of self- and other-perceptions.

The research presented here provides a framework for examining such questions, and provides preliminary answers to these questions. A great deal more research is necessary to refine our understanding of self- and other-knowledge, but the results presented here provide a solid starting point. Furthermore, the methods used in my studies provide proof that accuracy research is possible, and much can be learned by using designs that combine multiple methods. It is my hope that researchers will take advantage of recent technological and methodological advances and incorporate informant measures and behavioral measures into much of their personality research. If all personality studies included self, informant, and behavioral measures, our understanding of self- and other-knowledge would grow immensely. As described above, the applications of such information are numerous.

## CONCLUSION

The person looks very different when examined from the inside than from the outside. People's self-perceptions are more negative than others' perceptions of them, they are more aware of their negative traits than their positive traits, and they fail to notice a substantial number of their own characteristics. From the outside, people seem to

possess many desirable attributes. Their behavior is fairly predictable, and our knowledge of their observable personality is quite good. Although these discrepancies between the person from the inside and the person from the outside can be baffling, they can also be very informative. The two perspectives often complement each other – one filling in the gaps of the other. Furthermore, even when both perspectives are accurate, they are often accurate in different ways such that taking both into account deepens our understanding of the person. Neither perspective is completely correct, both provide different pieces of the puzzle that makes up the entire person.

# Appendix A

**ACT  [SELF-REPORT VERSION]**

Compared to other people, how much do you do the following activities? Select a number beneath each activity.

1 – Much less than the average person
2
3
4 – About as much as the average person
5
6
7 – Much more than the average person

1. Spend time by yourself
much less than average   1   2   3   4   5   6   7   much more than average
2. Spend time with others
much less than average   1   2   3   4   5   6   7   much more than average
3. Talk on the phone
much less than average   1   2   3   4   5   6   7   much more than average
4. Talk with someone one-on-one
much less than average   1   2   3   4   5   6   7   much more than average
5. Talk in with people in groups (with more than just one other person)
much less than average   1   2   3   4   5   6   7   much more than average
6. Talk with people of your own sex
much less than average   1   2   3   4   5   6   7   much more than average
7. Talk with people of the other sex
much less than average   1   2   3   4   5   6   7   much more than average
8. Laugh
much less than average   1   2   3   4   5   6   7   much more than average
9. Sing or whistle
much less than average   1   2   3   4   5   6   7   much more than average
10. Cry
much less than average   1   2   3   4   5   6   7   much more than average
11. Argue or fight
much less than average   1   2   3   4   5   6   7   much more than average
12. Listen to the radio or music
much less than average   1   2   3   4   5   6   7   much more than average
13. Watch TV
much less than average   1   2   3   4   5   6   7   much more than average
14. Spend time on the computer
much less than average   1   2   3   4   5   6   7   much more than average

15. Read
much less than average   1  2  3  4  5  6  7   much more than average
16. Work (at a job)
much less than average   1  2  3  4  5  6  7   much more than average
17. Spend time eating (not the amount eaten, but the time)
much less than average   1  2  3  4  5  6  7   much more than average
18. Attend class
much less than average   1  2  3  4  5  6  7   much more than average
19. Spend time doing entertaining things (e.g., going to the movies, to a sporting event, playing arcade games)
much less than average   1  2  3  4  5  6  7   much more than average
20. Sleep
much less than average   1  2  3  4  5  6  7   much more than average
21. Spend time in a house or apartment (any, not just your own)
much less than average   1  2  3  4  5  6  7   much more than average
22. Spend time outside
much less than average   1  2  3  4  5  6  7   much more than average
23. Spend time in a car or bus
much less than average   1  2  3  4  5  6  7   much more than average
24. Go to coffee shops, bars, or restaurants
much less than average   1  2  3  4  5  6  7   much more than average
25. Exercise or play sports
much less than average   1  2  3  4  5  6  7   much more than average

# Appendix B

## ROUND-ROBIN RATING FORM

For each personality trait below, rate how well the trait describes each person in your group (including yourself) by writing their letter above a number along the spectrum from "Not at all" to "Extremely". Rate each person compared to the average UT student.
For example, if the trait is "extraverted" and you think that person A is extremely extraverted compared to most UT students, you would write the letter A above the number 14 or 15, then go on to the next trait and continue rating person A. When you are done with person A, start over from the beginning and rate the next person. Rate yourself last. To show that you have read these instructions, please cross out the last word of this sentence.
Circle a different number for each person (two people cannot be on the same circle) and try to use the entire spectrum whenever appropriate.

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Extraverted, enthusiastic | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 2. Critical, quarrelsome | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 3. Dependable, self-disciplined | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 4. Anxious, easily upset | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 5. Open to new experiences, complex | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 6. Reserved, quiet | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 7. Sympathetic, warm | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 8. Disorganized, careless | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 9. Calm, emotionally stable | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 10. Conventional, uncreative | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |
| 11. Happy, satisfied with life | Not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | Extremely |

| | Not at all | | | | | | | | | | | | | | Extremely |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 12. Intelligent | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 13. Has strong math skills | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 14. Has strong verbal skills | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 15. Physically attractive | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 16. Has an attractive face | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 17. Has an attractive body | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 18. Lonely | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 19. Has high self-esteem | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 20. Is a genuinely dependable and responsible person | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 21. Assertive | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 22. Tends to dominate group discussions | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 23. Impulsive | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 24. Has a strong need to be around others, doesn't like being alone | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 25. Thinks and associates ideas in unusual ways, has unconventional thought processes | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 26. Arrogant, thinks too much of him/herself | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 27. Politically liberal | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |

28. Is a good leader   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

29. Good at public speaking   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

30. Likeable   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

31. Depressed   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

32. Exaggerates his/her skills   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

33. Power-oriented, values power in self and others   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

34. Likes to be the center of attention   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

35. Pays attention to detail   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

36. Tends to like others   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

37. Tends to be liked by others   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

38. Honest   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

39. Funny   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

40. Has a strong drive to achieve, is motivated to do well   Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

LIKING

On a scale of 1 to 15, how much do you like each person in your group?
(Place each person's letter above a number, except for your own letter. Do not use the same number twice.)

Not at all   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Extremely

# Appendix C

**INSTRUCTIONS FOR THE LEADERLESS GROUP DISCUSSION**

"In this activity, you will all be representatives from different departments of the same company. You are all on your organization's Compensation Committee. Five employees have been recommended for a merit bonus by their supervisors. You will each be representing one candidate from your department.

While you would like to grant substantial bonuses to all the candidates, the profits of the organization will not permit it. There is only $18,500 in merit bonus funds available.

In your packet you will find information about your candidate, and a little bit of information on each of the other candidates. You are under strong pressure from your department to get as much money for this candidate as possible. Your tasks during the committee discussion are to present a strong argument for your candidate and at the same time to help the committee decide the best allocation of the available funds.

The committee must reach a written decision in 10 minutes or no one receives a bonus. This is the last meeting of the year.

I will now give you a few moments to read over your packet, which contains detailed information about your candidate and a brief overview of the other candidates."

Start stopwatch, countdown from 1 min.

"Ok, before we begin, does anyone have any questions?

Ok, here is the form you must complete in the next ten minutes. At the beginning of the meeting, each committee member must give a 30-second presentation concerning his or her candidate. You must reach an agreement and write down your agreement on this form before the ten minutes are up. I will give you a warning when you have 1 minute left."

Place decision form and pen in center of table, equidistant from all participants.

# Appendix D

"Basically, the idea is just to write a complete story about each picture - an imaginative story with a beginning, a middle, and an end. Try to portray who the people in each picture might be, what they are feeling, thinking, and wishing for. Try to tell what led to the situation depicted in each picture and how everything will turn out in the end. Beneath each picture there are some guiding questions — these should be used as guides to writing your story. You do not need to answer them specifically. Look at the picture for a few seconds first, then turn the page and write whatever story comes to your mind. Don't worry about grammar, spelling, or punctuation — they are of no concern here. And if you need more space, use the back of the page. You will have about four minutes for each story — I will let you know when you should be finishing each one and moving on to the next. Does anybody have any questions? Remember, all of your stories are completely confidential. Do not go on to the next picture until I tell you to. Ok, go ahead and flip to the first picture and begin your story."

# References

Agnew, C. R., Loving, T. J., & Drigotas, S. M. (2001). Substituting the forest for the trees: Social networks and the prediction of romantic relationship state and fate. *Journal of Personality and Social Psychology, 81,* 1042-1047.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 256-274.

Ames, D. R., Rose, P., & Anderson, C. P. (in press). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality.*

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9,* 9-30.

Bass, B. M. (1954). The leaderless group discussion as a leadership evaluation instrument. *Personnel Psychology, 7,* 470-477.

Beck, A. T., Rial, W. Y., & Rickels, K. (1974). Short form of the Depression Inventory: Cross-validation. *Psychological Reports, 34,* 1184-1186.

Bem, D. J. (1967). Self-perception: An alternative explanation of cognitive dissonance phenomena. *Psychological Review, 74,* 183-200.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81,*

506-520.

Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person
perception accuracy: Another look at self-other agreement. *Personality and
Social Psychology Bulletin, 20,* 367-378.

Berry, D. S., & Finch Wero J. L. (1993). Accuracy of face perception: A view from
ecological psychology. *Journal of Personality, 61,* 497-503.

Bleske, A. L., & Buss, D. M. (2000). Can men and women be just friends? *Personal
Relationships, 7,* 131-151.

Block, J. (1989). Critique of the act frequency approach to personality. *Journal of
Personality and Social Psychology, 56,* 234-245.

Borkenau, P., & Lieber, A. (1993). Convergence of stranger ratings of personality and
intelligence with self-ratings, partner ratings, and measured intelligence. *Journal
of Personality and Social Psychology, 65,* 546-553.

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin
slices of behavior as cues of personality and intelligence. *Journal of Personality
and Social Psychology, 86,* 599-614.

Botwin, M. D., Buss, D. M., & Shackelford, T. K. (1997). Personality and mate
preferences: Five factors in mate selection and marital satisfaction. *Journal of
Personality, 65,* 107-136.

Branje, S. J. T., van Aken, M. A. G., van Lieshout, C. F. M., & Mathijssen, J. J. J. P.
(2003). Personality judgment in adolescents' families: The perceiver, the target,
their relationship, and the family. *Journal of Personality, 71,* 49-81.

Buss, D. M. (1987). Selection, evocation, manipulation. *Journal of Personality and*

*Social Psychology, 53,* 1214-1221.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses

tested in 37 cultures. *Behavioral and Brain Sciences, 12,* 1-49.

Buss, D. M. (1996). Social adaptation and five major factors of personality. In J. S.

Wiggins (Ed.), *The Five-Factor Model of Personality: Theoretical Perspectives*

(pp. 180-207). New York, NY: Guilford Press.

Buss, D. M., Abbott, M., Angleitner, A., Biaggio, A., Blanco-Wlasenor, A., & Bruchon-

Schweitzer, M. (1990). International preferences in selecting mates: A study of 37

cultures. *Journal of Cross-Cultural Psychology. 21,* 5-47.

Buss, D. M., & Craik, K. H. (1985). Why *not* measure that trait? Alternative criteria for

identifying important dispositions. *Journal of Personality and Social Psychology,*

*48,* 934-946.

Buss, D. M., Shackelford, T. K., Kirkpatrick, L. A., & Larsen, R. J. (2001). A half

century of mate preferences: The cultural evolution of values. *Journal of*

*Marriage and the Family, 63,* 491-503.

Buunk, B. P., Dijkstra, P., Fetchenhauer, D., Kenrick, D. T. (2002). Age and gender

differences in mate selection criteria for various involvement levels. *Personal*

*Relationships, 9,* 271-278.

Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality

tests: A peer-rating study. *Journal of Personality and Social Psychology, 43,*

1254-1269.

Colvin, C. R. (1993). "Judgable" people: Personality, behavior, and competing

explanations. *Journal of Personality and Social Psychology, 64,* 861-873.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68,* 1152-1162.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual.* Odessa, FL: Psychological Assessment Resources.

Cronbach, L. J. (1955). Processes affecting scores on "understanding others" and "assumed similarity." *Psychological Bulletin, 52,* 177-193.

Danner, D. D., Snowdon, D. A., & Friesen, W. V. (2001). Positive emotions in early life and longevity: Findings from the nun study. *Journal of Personality and Social Psychology, 80,* 804-813.

Davidson, K. W. (1993). Suppression and repression in discrepant self-other ratings: Relations with thought control and cardiovascular reactivity. *Journal of Personality,61,* 669-691.

Dunning, D. (2005). *Self-Insight: Roadblock and Detours on the Path to Knowing Oneself.* New York: Psychology Press.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83,* 1441-1455.

Epstein, S. (1983). Aggregation and beyond: Some basic issues on the prediction of behavior. *Journal of Personality, 51,* 360-392.

Fiedler, E. R., Oltmanns, T. F., & Turkheimer, E. (2004). Traits associated with personality disorders and adjustment to military life: Predictive validity of self and peer reports. *Military Medicine, 169,* 207-211.

Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density-distributions of states. *Journal of Personality and Social Psychology, 80,* 1011-1027.

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Tomlinson-Keasey, C., Martin, L. R., Wingard, D. L., & Criqui, M. H. (1995a). Childhood conscientiousness and longevity: Health behaviors and cause of death. *Journal of Personality and Social Psychology, 68,* 696-703.

Friedman, H. S., Tucker, J. S., Schwartz, J. E., Tomlinson-Keasey, C., Martin, L. R., Wingard, D. L., & Criqui, M. H. (1995b). Psychosocial and behavioral predictors of longevity: The aging and death of the "Termites." *American Psychologist, 50,* 69-78.

Friedman, R. S., & Förster, J. (2002). The influence of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology, 38,* 41-55.

Fuhrman, R. W., & Funder, D. C. (1995). Convergence between self and peer in response-time processing of self-relevant information. *Journal of Personality and Social Psychology, 69,* 961-974.

Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality, 48,* 473-493.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception.* San Diego, CA: Academic Press.

Funder, D. C. (2001). Personality. *Annual Review of Psychology, 52,* 197-221.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship,

agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55,* 149-158.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409-418.

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintance. *Journal of Personality and Social Psychology, 69,* 656-672.

Gill, M. J., & Swann, W. B. Jr. (2004). On what it means to know someone: A matter of pragmatics. *Journal of Personality and Social Psychology, 86,* 405-418.

Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology, 74,* 1337-7349.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82,* 379-398.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality, 37,* 504-528.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist, 59,* 93-104.

Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity:

Implications for self-peer agreement in personality judgment. *Journal of Personality and Social Psychology, 72,* 664-677.

Heller, D., & Watson, D. (2005). The dynamic spillover of satisfaction between work and marriage: The role of time and mood. *Journal of Applied Psychology, 90,* 1273-1279.

Heller, D., Watson, D., & Ilies, R. (2004). The role of person versus situation in life satisfaction: A critical examination. *Psychological Bulletin, 130,* 574-600.

Heller, D., Watson, D., Komar, J., Min, J., & Perunovic, W. O. E. (in press). Contextualized personality: Traditional and new assessment procedures. *Journal of Personality.*

Hogan, R. (1998). Reinventing personality. *Journal of Social and Clinical Psychology, 17,* 1-10.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*, 100-112.

Hogan, R., & Hogan, J. (2001). Assessing leadership: A view from the dark side. *International Journal of Selection and Assessment, 9,* 40-51.

Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality, 8,* 149-162.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the

unique perspective of the self. *Journal of Personality, 61,* 521-551.

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual

differences in self-enhancement and the role of narcissism. *Journal of Personality*

*and Social Psychology, 66*, 206-219.

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: Guilford Press.

Jones, E. E., & Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions*

*of the causes of behavior.* Morristown, NJ: General Learning Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York:

Guilford Press.

Kenrick, D. T., Groth, G. E., Trost, M. R., & Sadalla, E. K. (1993). Integrating

evolutionary and social exchange perspectives on relationships: Effects of gender,

self-appraisal, and involvement level on mate selection criteria. *Journal of*

*Personality and Social Psychology, 64,* 951-969.

Kirschbaum, C., Pirke, K. M., Hellhammer, D. H. (1993). The 'Trier Social Stress Test':

A tool for investigating psychobiological stress responses in a laboratory setting.

*Neuropsychobiology, 28,* 76-81.

Kolar, D. W., Funder, D. C., & Colvin, C. R. (1996). Comparing the accuracy of

personality judgments by the self and knowledgeable others. *Journal of*

*Personality, 64,* 311-337.

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004).

Reconceptualizing individual differences in self-enhancement bias: An

interpersonal approach. *Psychological Review, 111,* 94-110.

Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero-
acquaintance: A social relations analysis. *Journal of Personality and Social
Psychology, 65,* 1178-1187.

Lopes, P. N., Salovey, P., & Straus, R. (2003). Emotional intelligence, personality, and
the perceived quality of social relationships. *Personality and Individual
Differences, 35,* 641-658.

Luft, J. and Ingham, H. (1955) The Johari window, a graphic model of interpersonal
awareness. *Proceedings of the Western Training Laboratory in Group
Development.* Los Angeles: UCLA.

Marshall, G. N., Wortman, C. B., Vickers, R. R., Jr., Kusulas, J. W.,  & Hervig, L. K.
(1994). The Five-Factor Model of personality as a framework for personality-
health research. *Journal of Personality and Social Psychology, 67,* 278-286.

Martin, P. & Bateson, P. (1993). *Measuring Behavior: An Introductory Guide.*
Cambridge, UK: Cambridge University Press.

Martin, L. R., Friedman, H. S., Tucker, J. S., Schwartz, J. E., Criqui, M. H., Wingard, D.
L., & Tomlinson-Keasey, C. (1995). An archival prospective study of mental
health and longevity. *Health Psychology, 14,* 381-387.

McAdams, D. P. (1996). Personality, modernity, and the storied self: A contemporary
framework for studying persons. *Psychological Inquiry, 7,* 295-321.

McClelland, D. C. (1975). *Power: The Inner Experience.* New York: Irvington
Publishers.

McClelland, D. C., & Steele, R. S. (1972). *Motivational Workshops*. New York: General Learning Press.

McCrae, R. R., & Costa, P. T., Jr., (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology, 51,* 882-888.

McCrae, R. R., & Costa, P. T., Jr., (1999). A five-factor theory of personality. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 139-153). New York: Guilford Press.

Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology, 84,* 857-870.

Mehl, M. R., Pennebaker, J. W., Crow, M. D., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers, 33,* 517-523.

Mischel, W., & Shoda, Y. (1999). Integrating dispositions and processing dynamics within a unified theory of personality: The cognitive-affective personality system. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality theory and research* (pp. 139-153). New York: Guilford Press.

Nisbett, R. E. & Ross, L. D. (1980). Human inference: Strategies and shortcomings of social judgment. New York: Prentice-Hall.

Oltmanns, T. F., Gleason, M. E. J., Klonsky, E. D., & Turkheimer, E. (2005). Meta-perception for pathological personality traits: Do we know when others think that we are difficult? *Consciousness and Cognition: An International Journal, 14,*

739-751.

Oltmanns, T. F., Turkheimer, E., & Strauss, M. E. (1998). Peer assessment of personality traits and pathology in female college students. *Assessment, 5,* 53-65.

Ones, D. S., & Viswesvaran, C. (2001). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. *International Journal of Selection and Assessment, 9,* 31-39.

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74,* 1197-1208.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63,* 816-824.

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,* 890-904.

Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deception styles with basic traits and motives. *Journal of Personality, 66,* 1025-1060.

Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy IQ tests? *Journal of Personality, 66,* 525-554.

Paulhus, D. L., & Vazire, S. (in press). The self-report method in personality assessment. In R. W. Robins, R. C. Fraley, and R. Krueger (Eds.), *Handbook of Research*

*Methods in Personality Psychology.* New York: Guilford Press.

Pelham, B. W., & Swann, W. B. (1994). The juncture of intrapersonal and interpersonal knowledge: Self-certainty and interpersonal congruence. *Personality and Social Psychology Bulletin, 20,* 349-357.

Pickhardt, D., Vazire, S., & Gosling, S. D. (2003, February). *What Makes a Good Judge? An Exploration of Accuracy across Traits and Contexts.* Poster presented at the annual meetings of Society for Personality and Social Psychology, Burbank, CA.

Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54,* 890-902.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84,* 1236-1256.

Reynolds, D. J. Jr., & Gifford, R. (2001). The sounds and sights of intelligence: A lens model channel analysis. *Personality and Social Psychology Bulletin, 27,* 187-200.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychology, 80,* 340-352.

Robins, R. W., Caspi, A., Moffitt, T. E. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality, 70,* 925-964.

Robins, R. W., & John, O. P. (1997). Effects of visual perspective and narcissism on self-perception: Is seeing believing? *Psychological Science, 8,* 37-42.

Robins, R. W., Noftle, E. E., Trzesniewski, K. H., & Roberts, B. W. (2005). Do people know how their personality has changed? Correlates of perceived and actual personality change in young adulthood. *Journal of Personality, 73,* 489-521.

Rohrbaugh, M. J., Shoham, V., Coyne, J. C., Cranford, J. A., Sonnega, J. S., & Nicklas, J. M. (2004). Beyond the 'self' in self-efficacy: Spouse confidence predicts patient survival following heart failure. *Journal of Family Psychology, 18,* 184-193.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54,* 93-105.

Slatcher, R. B., & Vazire, S. (2006). It's not just who you are, it's how you act: Influences of global and contextualized personality traits on relationship satisfaction. Manuscript submitted for publication.

Slatcher, R. B., Vazire, S., & Pennebaker, J. W. (2006). A view from the inside: Predicting relationship stability from couples' everyday interactions. Manuscript submitted for publication.

Slate.com's "Dear Prudence" column (n.d.). Retrieved February 10th, 2005, from http://www.slate.com/id/2112478/

Smith, C. P. (Ed.). (1992). *Motivation and personality: Handbook of thematic content analysis*. New York: Cambridge University Press.

Spain, J. S., Eaton, L. G., & Funder, D. C. (2000). Perspectives on personality: The relative accuracy of self versus others for the prediction of emotion and behavior. *Journal of Personality, 68,* 837-867.

Sprecher, S., & Regan, P. C. (2002). Liking some things (in some people) more than others: Partner preferences in romantic relationships and friendships. *Journal of Social and Personal Relationships, 19,* 463-481.

Swann, W. B. Jr., (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91,* 457-477.

Swann, W. B., & Read, S. J. (1981). Self-verification processes: How we sustain our self-conceptions. *Journal of Experimental Social Psychology, 17,* 351-372.

Swann, W. B., & Rentfrow, P. J. (2001). Blirtatiousness: Cognitive, behavioral, and physiological consequences of rapid responding. *Journal of Personality and Social Psychology, 81,* 1160-1175.

Taylor, S. E, & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103, 193-210.*

Tucker, J. S., Friedman, H. S., Wingard, D. L, & Schwartz, J. E. (1996). Marital history at midlife as a predictor of longevity: Alternative explanations to the protective effect of marriage. *Health Psychology, 15,* 94-101.

Vazire, S. (in press-a). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality.*

Vazire, S. (in press-b). Looking glass self. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of Social Psychology.* Thousand Oaks, CA: Sage publications.

Vazire, S., & Funder, D. C. (in press). Impulsivity and the self-defeating behavior of narcissists. *Personality and Social Psychology Review.*

Vazire, S., & Gosling, S. D. (2003, October). *Accuracy and Bias in Self-Perception: A*

*New Approach to an Old Question.* Poster presented at the annual International
Positive Psychology Summit, Washington, D. C.

Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on
personal websites. *Journal of Personality and Social Psychology, 87,* 123-132.

Vazire, S., Gosling, S. D., Dickey, A. S., & Schapiro, S. J. (in press). Measuring
personality in non-human animals. In R. W. Robins, R. C. Fraley, and R. Krueger
(Eds.), *Handbook of Research Methods in Personality Psychology.* New York:
Guilford Press.

Vazire, S., & Robins, R. W. (2004). Beyond the justification hypothesis: A broader
theory for the evolution of self-consciousness. *Journal of Clinical Psychology, 60,*
1271-1273.

Wagerman, S. A., Greve, L. A., Wright, L. D., & Funder, D. C. (2004, January). *Making
the Grade: Personality Correlates of Academic Achievement.* Poster presented at
the annual meeting of the Society for Personality and Social Psychology, Austin,
TX.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of a Brief
Measure of Positive and Negative Affect. *Journal of Personality and Social
Psychology, 54*, 1063-1070.

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and
affectivity: The role of acquaintanceship, trait visibility, and assumed similarity.
*Journal of Personality and Social Psychology, 78,* 546-558.

Wilson, T.D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious.* Cambridge, MA: Belknap Press/Harvard University Press.

Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, values, and potential for improvement. *Annual Review of Psychology, 55,* 493-518.

Wonderlic, E. F. (1983). *Wonderlic Personnel Test Manual.* Northfield, IL: E. F. Wonderlic & Associates.

Wood, D., & Harms, P. D. (2005). *Contributions of role identities and role information in understanding contextual effects in person perception.* Unpublished manuscript.

Wood, D., & Roberts, B. W. (in press). Cross-sectional and longitudinal tests of the personality and role identity structural model (PRISM). *Journal of Personality.*

# Vita

Simine Vazire was born in Grenoble, France on February 28th, 1980, the daughter of Monique Vazire and Hamid Vazire, and the younger sister of Navid Vazire. She grew up mostly in Palo Alto, California and attended Henry M. Gunn High School. She then attended Carleton College in Northfield, Minnesota. After completing her B.A. in psychology with a concentration in women's studies in 2000, she moved to Minneapolis and worked for a year at the Center for Twin and Family Studies in the Psychology Department at the University of Minnesota, and for the Minnesota Senate's office of Senate Counsel and Research. In the fall of 2001, she began graduate school at UT Austin under the guidance of Sam Gosling. In the 2003-2004 academic year she conducted research with David Funder at the University of California, Riverside in the fall semester and with Del Paulhus at the University of British Columbia in the spring semester before returning to UT Austin in the fall of 2004. In 2005, she was awarded a William S. Livingston University Continuing Fellowship and a dissertation fellowship from the Council of Graduate Departments of Psychology. After graduating in May of 2006, Simine will be a visiting assistant professor at the University of Virginia in the fall of 2006 and a post-doc at UT Austin in the spring of 2007 before assuming a position as an assistant professor in the Psychology Department at Washington University in St. Louis in the summer of 2007.

Permanent address:     3477 Kenneth Dr., Palo Alto, CA 94303

This dissertation was typed by the author.