AN ABSTRACT OF THE THESIS OF

Jill A. Brown for the degree of Honors Baccalaureate of Science in Psychology presented

on May 22, 2009. Title: Perceiving Big Five Traits Accurately at True Zero

Acquaintance.

Abstract Approved: _____

Frank J. Bernieri

This research focused on the accuracy of trait judgments at zero acquaintance and

after five minutes of interaction. Thirty-five undergraduates judged the neuroticism,

extraversion, openness, agreeableness, and conscientiousness of four to six others whom

they did not know. Participants then interacted in one-on-one conversations with each

target for five minutes. It was hypothesized that extraversion and conscientiousness

would be more accurately judged at zero acquaintance than neuroticism, openness, or

agreeableness. It was also predicted that all traits would be more accurately judged after

five minutes and that the NEO PI-R, a validated personality assessment, would be a more

accurate measure of the criterion than a brief self report consisting of ten items. Contrary

to expectations, significant accuracy was found at zero acquaintance for only the trait of

neuroticism. Extraversion accuracy did not appear until after the five minute

conversation.

Perceiving Big Five Traits Accurately at True Zero Acquaintance

by

Jill A. Brown

A PROJECT

Submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in Psychology (Honors Scholar)

Presented May 22, 2009

Commencement June 13, 2009

ACKNOWLEDGEMENTS

project of <u>Jill A. Brown</u> presented on

<u>May 22, 2009</u>.

APPROVED:

_____

Mentor, representing Psychology

_____

Committee Member, representing Psychology

_____

Committee Member, representing Psychology

_____

Chair, Department of Psychology

_____

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

_____

Jill A. Brown, Author

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

This thesis is dedicated to those before me who know this process well

and to those after me who are just beginning the journey.

Perceiving Big Five Traits Accurately at True Zero Acquaintance

Many situations in life call upon us to impress people. When going for that first nerve-racking job interview we know to put our best foot forward because at some point we were told those first impressions are the most important. We know that the interviewer is going to be evaluating whether we are the right candidate for the job. This evaluation may include criteria ranging from charisma and enthusiasm to how we are dressed to how well our personality will mesh with current employees. These factors, and others, may be important to consider when presenting oneself to another for the first time.

Whether we realize it or not, we form impressions of others immediately (DePaulo, Kenny, Hoover, Webb, & Oliver, 1987). When meeting someone new, we make inferences about each others' personality (Kenny, 1994). First impressions also influence the outcomes of further interactions (Ambady, Hallahan, & Rosenthal, 1995). For example, when teachers were told certain students in their classrooms were "spurters" with exceptional talents, they treated them as such and gave them special attention when in fact they were randomly chosen (Rosenthal & Jacobson, 1968). These children showed greater gains in IQ than other children in the classroom simply because of the belief held by the teachers and their resulting behavior. Thus, what we think about a person in the very beginning of the relationship can greatly influence our lasting perception of that individual (Prickett, Gada-Jain, & Bernieri, 2000). Oftentimes appearance is the only information available to shape initial judgments, such as those of personality (Shevlin, Walker, Davies, Banyard, & Lewis, 2003).

Much research on impression formation has been done with unacquainted individuals (e.g. Ambady et al., 1995). This particular research paradigm, known as zero

acquaintance, looks at initial interactions and judgments of first impressions (Albright, Kenny & Malloy, 1988). Zero acquaintance is defined as a situation where judges have no prior interaction with targets (Albright et al., 1988). This first impression can involve a variety of facets of personality (e.g. affect and attitudes), but in much of the zero acquaintance literature the big five personality traits of neuroticism, extraversion, openness, agreeableness, and conscientiousness have been the primary focus (e.g. Albright et al, 1988; Borkenau & Liebler, 1992).

Typical zero acquaintance situations involve two or more individuals rating each other on certain aspects of personality. The Albright et al. (1988) study asked participants in a classroom to silently form groups of five individuals with others who they had no previous acquaintance. A round-robin design was employed where each participant served as both a judge and a target. They found that for extraversion and conscientiousness a significant amount of the variance in ratings was due to the target, meaning that judges were able to discriminate between the targets on these facets of personality.

Another zero acquaintance study (Borkenau & Liebler, 1992) used video footage of targets. Borkenau and Liebler (1992) reported that judges were accurate at detecting extraversion, $r = 0.33$, $p < 0.01$ and conscientiousness, $r = 0.32$, $p < 0.01$ but not neuroticism, $r = 0.06$, $p > 0.05$ at zero acquaintance. Kenny, Horner, Kashy, & Chu (1992) found that judges agree on a target's extraversion and to a lesser extent, conscientiousness. Levesque and Kenny (1993) also found significant judge agreement for ratings of extraversion but not for the other traits. Thus, the literature to date suggests

that extraversion and, to a lesser extent, conscientiousness are the only two traits accurately judged at zero acquaintance.

Borkenau and Liebler (1992) found that extraversion accuracy at zero acquaintance had a strong relationship with certain physical attributes making it more visible than other traits at this point in time. Other researchers have found that perceiver agreement in the extraversion of others was increased by physical attractiveness (Albright et al., 1988; Kenny, Horner, Kashy & Chu, 1992). Because appearance is the main source of information for zero acquaintance judgments (Shevlin et al., 2003), accuracy would not be expected for any trait where the cues for that trait are not manifested in appearance. Visibility of emotional stability, also known as neuroticism, may be variable because the outward manifestation of this trait is limited unless provoked by the anxiety of the situation (Kenny et al., 1992). Neuroticism was least accurately judged using physical cues at zero acquaintance (Borkenau & Liebler, 1992), suggesting it is not a trait whose judgment can be based merely on appearance.

From these studies, it was predicted that judges would be able to accurately assess extraversion and conscientiousness at zero acquaintance. However, neuroticism and agreeableness accuracy were not predicted to be judged accurately at zero acquaintance.

Level of Acquaintance and Accuracy

A personality judgment is subject to the influence of outside factors such as how long one person has known another (Bernieri, Zuckerman, Koestner, & Rosenthal, 1994). Theoretically the longer you have known someone, the more time you have had to interact and gather relevant information. This information can be used to make a

judgment about certain aspects of their personality and form an overall impression. Bernieri et al. (1994) reported that roommates that had known each other for a longer period of time had more agreement in ratings of neuroticism and agreeableness than those who had known each other for a short time, suggesting that neuroticism, for example, may not have been judged accurately initially. Extraversion and openness showed no increase in the accuracy coefficients, indicating that individuals may have been accurate at zero acquaintance and their ratings did not change. The most significant decrease in accuracy due to acquaintance level was found for ratings of conscientiousness (Bernieri et al., 1994), which seems to contradict other research (Albright et al., 1988).

Kenny et al. (1992) found that targets who were judged to be extraverted at zero acquaintance were also rated as extraverted after one-on-one conversations. Blackman & Funder (1996) found that more information (i.e. increasingly longer videotaped scenarios) about a target leads to greater self-other agreement about personality items on the California Q-set. This suggests that at the minimum, judgment accuracy generally should increase as a result of more information, provided the information is relevant.

McLarney-Vesotski, Bernieri, and Rempala (2006) designed a study to elicit relevant cues for each personality trait in a series of contexts. The sequence was shown to judges and in total depicted less than five minutes of behavior of each target, qualifying for a 'thin slice' of behavior (Ambady et al., 2000). Employing trait relevant situations enabled perceivers to accurately judge neuroticism and openness after less than five minutes of information. Their results showed that accuracy for other traits may be possible during a brief initial encounter.

Previous studies have given mixed results for the accuracy in judging particular traits over time (e.g. Bernieri et al., 1994; Borkenau & Liebler, 1992). The current study employed the zero acquaintance paradigm to assess first impression judgments and also introduced a five minute interaction manipulation: a period of one-on-one conversation between dyads. It was expected that all traits should yield an increase in judgment accuracy after five minutes of interaction. At this point in the relationship, judges would have had access to more relevant information on which to base their judgments. The main source of information at zero acquaintance is physical appearance; the types of clothes worn as well as how kempt an individual looks may both impact personality judgments. After interacting for five minutes, judges will have knowledge of the target's speaking style, eye contact, and enthusiasm in addition to any specific personal details the target elected to divulge. Therefore, it was hypothesized that judges would be more accurate at judging an individual's personality after five minutes of conversation than at zero acquaintance.

Methods of Accuracy Measurement

In order to assess accuracy of personality judgments, accuracy needs to be operationalized. In the beginning of person perception research, consensus scoring was often the way accuracy was determined (Albright et al., 1988). In this type of accuracy assessment, the agreement of all other observers is taken as accuracy. For example, if most everyone identified Suzie as extraverted, she was considered to be extraverted. In this case, it is not Suzie's self-reported extraversion or tested extraversion that is the correct answer, but how others *perceive* her extraversion. This type of scoring is often

used because if there is agreement between perceivers, then it is suggested that the judged attributes exist (Borkenau & Liebler, 1992).

Some researchers have used an established personality measure as their accuracy criterion, however, the Neuroticism-Extraversion-Openness Personality Inventory - Revised (NEO PI-R) (Costa & McCrae, 1992) is often too lengthy for research purposes, taking about 45 minutes to complete (Gosling, Rentfrow, & Swann Jr., 2003). Though often time consuming, this type of criterion may be preferred in the assessment of personality because it provides a more valid measure of the big five personality traits. Many studies, however, have used a personality accuracy criterion based on self-report data, which include the measures filled out by participants rating their own traits (Borkenau & Liebler, 1992). Presumably this is because a self-report measure is more similar to what judges are using to make trait ratings. Furthermore, self report scales that are matched to the scales used by others allows for the calculation of difference scores as a metric of accuracy (Cronbach, 1955). And, there is evidence to suggest that self ratings and scores on personality inventories are highly correlated (Gosling et al., 2003; Merrens, 1975). However, self ratings often carry their own confounds.

Self ratings of personality judgments may be subject to the social desirability bias (Dunnett, Koun, & Barber, 1981). That is, targets may not report where they actually feel they fall on a particular trait but instead bias their answers to be more socially desirable. For example, an individual who knows she is high on neuroticism may report a slightly lower score in order to better manage her impressions on others (Dicken, 1959). This bias could be problematic when using a brief, face valid measure of personality such as the

Ten-Item Personality Inventory (TIPI) (Gosling et al., 2003). Because there are only ten items on the TIPI, it may be apparent which facet of personality each item is referencing.

This measurement in effect would not be representative of the judge's ratings of another's personality, but their knowledge of what is socially desirable. An individual's accuracy for judging a particular trait may be increased by simply responding to items in the most socially desirable manner (Dunnett et al., 1981). Therefore, it is hypothesized that judges' ratings of targets' personality will be more accurate when target traits are measured by the more valid NEO PI-R than by a brief ten item self report.

Quantifying Accuracy

There are many ways to calculate the accuracy of a personality judgment. One method already mentioned is to use a difference score to determine each person's deviation from the criterion (Cronbach, 1955). A difference score would represent the degree of divergence from the accuracy criterion. Judges and targets may not use the rating scale in the same manner and therefore cannot be reliably compared (Sharon, Bartlett, 1969). For example, if participant A rates themselves as a seven on extraversion, the highest in the group, participant B may rate A's extraversion as a five while participant C rates A's extraversion as a six. Using a difference score, participant C would be more accurate than B. However, if both of their ratings for A were the highest in their group, they should be equally accurate (Bernieri et al., 1994).

Cronbach (1955) suggested a means of re-evaluating accuracy measurement based on the 'd' statistic which included breaking it down into several different components. Fortunately we do not need to use a difference score because a correlation

coefficient can also be used in the determination of accuracy (DePaulo et al., 1987). The correlation assesses how an individual is able to accurately identify the rank order of their group members on a particular trait. That is, they are able to distinguish the group members that are high on a trait from group members low on a trait. This type of accuracy assessment takes into account the different ways participants may use the rating scale. In other words, a correlation coefficient measures how well the judge's ratings of a sample of targets covaries with the criterion scores known for them.

Objective

The objective of the current study was to determine the accuracy of personality judgments over time. It was hypothesized that extraversion and conscientiousness would be more accurately judged at zero acquaintance than the traits of neuroticism, openness, or agreeableness. It was further predicted that all traits would show an increase in judgment accuracy after five minutes of interaction. Finally, it was hypothesized that judges' ratings of targets' personality will be more accurate when target traits are measured by the more valid NEO PI-R than by the TIPI, a brief ten item self report.

Method

Participants

Thirty-five students (13 males and 22 females) enrolled in a ten week long research practicum offered by the OSU Department of Psychology for which they received academic credit. The age range was from 19 to 42 with a mean at 22 years. The

course met three times a week. As part of this practicum they were assessed and received feedback on a variety of psychological measures and emotional intelligence tests. Participants engaged in a number of interpersonal activities and small group activities (e.g. games). Students earned a pass/fail grade for this course based solely on attendance.

Procedures

The first personality judgment was made before any interaction could take place. Prior to entering the meeting room on the first day of the term, participants were told that there would not be any talking for the day and they were not to communicate with others in the room. Participants were then invited into the classroom after being given a syllabus and a coded identification, which was in the form of a nametag with a letter of the alphabet as their identifier for the term. Participants were instructed to go to the desk and chair that matched their ID. Participants were randomly divided into groups ranging in size from five to seven members. The male to female ratio in each group was comparable. After all participants had arrived, instructions were given to the group. There was a check performed to ensure participants were not familiar with other group members. This was done as both a verbal exchange and as a paper-and-pencil task designed to assess the level of acquaintance between group members. Participants were then directed to turn their desks so that they were seated alphabetically in a circle facing inward (See Figure 1). All ratings were completed using a round-robin design where participants rated everyone in their group on all dimensions. Specifically, participant A would complete ratings for participants B, C, D, E, F, and G in that order. Participant B would rate participants A, C, D, E, F, G in that order, and so on. The seats were arranged

such that each individual began by judging the personality of the participant sitting to their right. By organizing it this way, there was not an instance where everyone in the group was looking at one individual at the same time. Groups were dismissed when everyone was finished.

Figure 1. Seating Arrangement
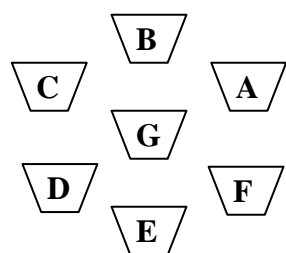
Pre-Zero Acquaintance
Arrangement
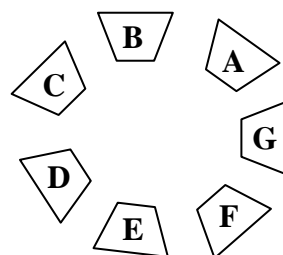
Round-Robin Arrangement

During the next two class meetings, each group member had the opportunity to converse with each other group member about anything they chose. The five minute interaction impressions were collected after these "getting to know you" one-on-one conversations lasting for five minutes. Each five minute conversation was followed by an impression rating packet similar to that used on the first day of class. The interactions were ordered such that one participant sat out and completed a separate measure while the other group members were talking and making their ratings. At this time participants rated themselves on the very same items used to judge others in their group. For example, during the first round of first impression judgments, participant A completed a separate packet containing various personality measures while participant B interacted with

participant C, D with E, and F with G. During each subsequent round, the participant with the following letter in the alphabet would sit out.

Several weeks later, all participants completed the NEO PI-R (Costa & McCrae, 1992) that provided criterion trait scores for the big five traits.

Materials

Neuroticism-Extraversion-Openness Personality Inventory - Revised (NEO PI-R) (Costa & McCrae, 1992).

The NEO PI-R is a 240 item assessment that is designed to measure comprehensively the Big Five personality constructs of Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. The test takes approximately thirty minutes to complete. The test uses a likert-type response format asking the individual the extent to which they agree or disagree with the given statements. Each of the NEO PI-R items corresponds to one of the big five personality traits. Each trait domain is divided into six facets, however only the five major trait domains are reported here. Representative items appear in Appendix A.

Personality Rating Scale.

Both the Zero-Acquaintance and Five Minute Interaction scales required participants to rate all others in their group on multiple attributes including current emotional state, personal relationships, predicted relationships, and stable traits.

Judgments of stable personality traits were made using an adaptation of Gosling, et al.'s Ten-Item Personality Inventory (TIPI) (2003) and served as the principle judgment

made at Zero-Acquaintance and Five-Minute Interaction (See Appendix B). The TIPI was designed to measure the Big Five personality constructs of Extraversion, Agreeableness, Neuroticism, Openness to Experience, and Conscientiousness. Each trait was represented by two items on a bipolar scale. The TIPI is presented in a likert-type format and takes less than five minutes to complete. A sample item is shown below:

not extraverted, not enthusiastic   1 2 3 4 5 6 7   extraverted, enthusiastic

Judges rated the degree to which the target was both extraverted and enthusiastic or not. A second item related to the trait was reverse-scored.

not reserved, not quiet   1 2 3 4 5 6 7   reserved, quiet (R)

To create the trait judgment score, the score from the first item and the reverse of the second item were averaged.

The current emotional state section asked how an individual appeared to be at that moment, asking the individual for example, to rate a group member on how nervous or relaxed the other appeared (See Appendix C). The personal relationships section included questions about family and romantic relations (See Appendix D) and the predicted relationships section asked the judge to guess how they will feel about the target at the end of the term (See Appendix E). The current state and relationship measures will not be included in this report because they are not relevant to the trait impression accuracy hypothesis being studied.

Accuracy Coefficient

Every participant generated a set of scores representing their judgments of five traits of every other individual in their group (i.e., neuroticism, extraversion, openness, agreeableness, and conscientiousness) and a set of scores which represented their own, self-assessed level. In addition, their domain scores on the NEO PI-R was also known.

To determine how accurate a participant was in judging their group members on a given trait, their judgments on that trait for their six group members were correlated with the accuracy criterion, in this case, the self-assessed personality scores from the NEO PI-R. For example, participant A's judgments of targets B through G's neuroticism were correlated with the targets' criterion score for neuroticism, which was generated by the NEO PI-R. A high positive correlation meant that participant A effectively identified the members in his or her group that were high and low on neuroticism. A negative accuracy correlation coefficient meant that a participant judged low neurotics as high and high neurotics as low.

In order to compare the NEO PI-R and TIPI rating scales, trait accuracy was also calculated using each participant's self-ratings on the TIPI as the judgment criterion.

Results

Table 1 displays the sample mean NEO PI-R trait scores that were used as accuracy criteria. The means for each trait are reported for each sex separately. The college population means identified by Costa and McCrae (1992) are also reported for comparison.

Table 1. *NEO PI-R[a] Means (n = 35)*

Trait	Population Mean[b]	Mean (STD)	Mean Male (n=13)	Mean Female (n=22)
Neuroticism	96	88 (25.93)	75	96
Extraversion	121	125 (18.94)	126	125
Openness	117	132 (18.31)	143	125
Agreeableness	114	117 (18.05)	114	119
Conscientiousness	115	115 (22.83)	118	113

[a] Costa & McCrae (1992). NEO Personality Inventory Revised.
[b] Population of college students (Costa & McCrae, 1992)

The means reported in this study mirror the population means reported in the NEO PI-R Professional Manual (Costa & McCrae, 1992) with the following exceptions. Males in this sample appeared more open than the population of college males previously reported. They also seemed less neurotic as the population, but this difference was not statistically significant[1].

Judgment Accuracy Coefficients: NEO PI-R as the Criterion

Judgments of each target were correlated with the each targets' score from the NEO PI-R. This meant that each judge generated five accuracy correlations; one for each trait they attempted to judge. These correlations were then transformed into Fisher zs.

These transformations standardized the correlations making them suitable to analyze as accuracy coefficients[2].

It was predicted that extraversion and conscientiousness would be judged more accurately at zero acquaintance than the traits of neuroticism, openness, and agreeableness. Table 2 shows that using the NEO PI-R as accuracy criteria, participants were not accurate at judging extraversion at zero acquaintance, $t_{(33)} = -0.15$, $p = 0.88$[3]. The lack of extraversion accuracy is contrary to most zero-acquaintance literature (Albright et al., 1988; Levesque & Kenny, 1993). Surprisingly, neuroticism was judged accurately at zero acquaintance, $t_{(33)} = 4.55$, $p < 0.01$. Neuroticism is not a trait normally reported to be accurately judged at zero acquaintance (Albright et al., 1988; Levesque & Kenny, 1993). Openness accuracy was approaching significance $t_{(33)} = 1.89$, $p < 0.07$, but neither agreeableness nor conscientiousness accuracy were significant, $t_{(33)} = 0.19$, $p < 0.85$, $t_{(33)} = -0.20$, $p = 0.84$, respectively. A one-way, five level (Trait) repeated measures contrast determined that neuroticism was judged significantly more accurately than the average of other traits at zero acquaintance within the NEO PI-R, $F_{(1, 32)} = 9.94$, $p < 0.01$.

Table 2. *Mean Judgment Accuracy Coefficients[a]: NEO PI-R as the accuracy criterion (n=35)*

Time	N	E	O	A	C
Zero Acq.	.42**	-.02	.22†	.03	-.01
Five Min.	.39**	.39**	.03	-.15	.16

†p[b]<0.10, *$p<0.05$, **$p<0.01$

[a] Coefficients are Fisher *z* transformations of Pearson *r* correlations.
[b] *p* values are report the significance against zero.

After participants had the chance to interact with each other for five minutes a slightly different pattern emerged. Judgments of extraversion became accurate, $t_{(33)} = 2.93$, $p < 0.01$. This level of accuracy is more consistent with zero acquaintance literature on extraversion judgments (Albright et al., 1988; Levesque & Kenny, 1993). Neuroticism continued to be judged accurately, $t_{(33)} = 3.30$, $p < 0.01$. Openness, agreeableness, or conscientiousness were not judged accurately even after five minutes of interaction, $t_{(33)} = 0.26$, $p = 0.79$, $t_{(33)} = -1.24$, $p = 0.22$, $t_{(33)} = 1.27$, $p = 0.21$, respectively. A one-way, five level (Trait) repeated measures contrast determined that accuracy in neuroticism judgments were approaching significance over the average of other traits after five minutes of interaction within the NEO PI-R, $F_{(1, 33)} = 3.84$, $p = 0.06$.

It was hypothesized that, across all five traits, judges would be more accurate after five minutes of interaction than they were at zero acquaintance. At this point, judges should have more trait-relevant information on which to base their judgments. In turn, judges would display increased accuracy in judging personality traits. Using a one-way (Time) ANOVA, it was determined that perceptions of extraversion became more accurate between zero acquaintance and five minute interaction, $F_{(1, 34)} = 6.84$, $p < 0.05$. No other traits showed a significant difference between the two time periods within the NEO PI-R.

Collapsing across length of acquaintance, 2×5 (Time $\times$ Trait) repeated measures contrast showed that neuroticism judgments were significantly more accurate than the mean of the other traits, $F_{(1, 32)} = 7.31$, $p < 0.02$. Agreeableness judgment accuracy was significantly lower than the mean of the other traits across length of acquaintance within the NEO PI-R, $F_{(1, 33)} = 4.19$, $p < 0.05$.

Judgment Accuracy Correlations: TIPI Self Ratings as the Criterion

Table 3 shows the mean self ratings for each trait as measured by the Ten Item Personality Inventory (TIPI) (Gosling et al., 2003). These scores were used as an alternative accuracy criterion with which to assess the observers' accuracy.

Table 3. *Mean Trait Self-Assessments (TIPI[a]) n=35*

Trait	Mean (STD)	Mean Male (n=13)	Mean Female (n=22)
Neuroticism[b]	3.04 (1.32)	2.62	3.30
Extraversion	4.67 (1.37)	4.69	4.65
Openness	5.39 (1.20)	5.15	5.52
Agreeableness	5.03 (1.05)	4.77	5.18
Conscientiousness	5.86 (1.03)	5.61	6.00

[a] TIPI self scores rated on a 7 point likert scale.
[b] Composite score is the mean of related items after reverse scoring.

Table 4 shows that TIPI criterion self ratings correlated significantly with criterion scores on the NEO PI-R in the present study for all traits except for Openness. Correlations for the other four traits were strong, ranging from .35 for agreeableness to .65 for conscientiousness. For comparison, Gosling et al. (2003) found significant correlations between each of the NEO PI-R and TIPI items, ranging from .56 for

openness and .68 for conscientiousness. These strong correlations replicate and confirm that the TIPI can be employed as a valid measure of the big five personality traits for impression accuracy studies.

Table 4. *Correlations between NEO PI-R[a] and TIPI[b] self*

	N	E	O	A	C
Brown, 2009	.65*	.50*	.13	.35*	.47*
Gosling et al., 2003	.66*	.65*	.56*	.59*	.68*

[a] NEO PI-R scores calculated by creating a composite variable combining each related item in the measure.
[b] TIPI self scores rated on a 7 point likert scale, composite is the mean of related items.

*p<0.05

As was true with the NEO PI-R criterion, it was hypothesized that at zero acquaintance judges would be accurate at rating extraversion and conscientiousness. Table 5 shows the judgment accuracy coefficients using the TIPI as accuracy criteria. Neither extraversion, $t_{(33)} = -0.19$, $p = 0.85$, nor neuroticism, $t_{(33)} = 1.34$, $p = 0.19$ judgments were significantly accurate at zero acquaintance. Conscientiousness judgments at zero acquaintance were approaching significance, but not in the predicted direction, $t_{(31)} = -1.73$, $p = 0.09$. The absence of neuroticism accuracy at zero acquaintance is a replication of previous studies (Kenny et al., 1992; Borkenau & Liebler, 1992), however, the lack of extraversion and conscientiousness accuracy is not consistent with the literature. However, accuracy was seen for agreeableness judgments at zero acquaintance,

$t_{(33)} = 3.39$, $p < 0.01$. Agreeableness accuracy at zero acquaintance is not traditionally

seen. A one-way, five level (Trait) repeated measures contrast determined that

agreeableness was judged significantly more accurately than the mean of other traits at

zero acquaintance within the TIPI, $F_{(1, 32)} = 9.06$, $p < 0.01$.

After five minutes of interaction it was predicted that judges would accurately rate

all traits. Judges were accurate at judging extraversion, $t_{(33)} = 4.06$, $p < .01$. Neuroticism

was also judged accurately at this point in the relationship, $t_{(33)} = 2.59$, $p < 0.02$.

However, there was no accuracy seen for openness, agreeableness, or conscientiousness

after five minutes of interaction, $t_{(33)} = 1.64$, $p = 0.11$, $t_{(33)} = -0.30$, $p = 0.76$, $t_{(32)} = -0.28$,

$p = 0.77$, respectively. A one-way, five level (Trait) repeated measures contrast

determined that extraversion was judged significantly more accurately than the mean of

other traits after five minutes of interaction within the TIPI, $F_{(1, 32)} = 5.76$, $p < 0.03$. This

replicates what is traditionally seen for zero acquaintance judgments (Albright et al.,

1988). It was also determined that agreeableness was judged significantly less accurately

than the mean of other traits after five minutes of interaction, $F_{(1, 32)} = 5.03$, $p < 0.04$.

Table 5. *Mean Judgment Accuracy Coefficients[a]: TIPI (n=35)*

Time	N	E	O	A	C
Zero Acq.	.19	-.03	.03	.40*	-.24†
Five Min.	.35*	.45**	.21	-.03	-.03

[a] Coefficients are Fisher z transformations of Pearson r correlations.

†$p<0.10$, *$p<0.05$, **$p<0.01$

Accuracy was hypothesized to increase for all five traits between zero acquaintance and five minute interaction. A one-way (Time) ANOVA was used to analyze this effect when using the TIPI as a criterion. There was not a significant difference between accuracy at zero acquaintance and at five minute interaction for neuroticism, $F_{(1, 34)} = 1.33$, $p = 0.26$. Accuracy for judging agreeableness significantly decreased after five minutes of interaction, $F_{(1, 32)} = 12.86$, $p < 0.02$. However, participants judged extraversion more accurately after five minutes of interaction than at zero acquaintance when employing the simple two-item self rating on the TIPI as the criterion, $F_{(1, 34)} = 8.26$, $p < 0.01$. The increase in accuracy between zero acquaintance and five minutes for judgments of extraversion was higher than the mean increases for the other traits, $F_{(1, 32)} = 4.65$, $p < 0.03$..

Judgment Accuracy Correlations: Overall Accuracy

It was hypothesized that regardless of criterion employed, judges would be more accurate at judging an individual's personality after five minutes of conversation than at zero acquaintance. This hypothesis was tested using a 2 × 2 × 5 (Criterion × Trait × Time) repeated measures ANOVA and approached significance, $F_{(1, 32)} = 3.53$, $p = 0.07$.

It was also predicted that at both zero acquaintance and five minute interaction, judges would be more accurate at reading a target's personality as measured by the NEO PI-R than when measured by self-ratings on the ten-item TIPI. It was determined by a 2 × 2 × 5 (Criterion × Time × Trait) repeated measures ANOVA that there was no difference,

$F_{(1, 32)} = 0.00$, $p = 0.96$. A contrast determined that neuroticism was the most accurately judged trait overall, collapsing across time and criterion, $F_{(1, 32)} = 8.83$, $p < 0.01$.

Discussion

Zero Acquaintance vs. Five Minute Interaction

Using the NEO PI-R as the criterion, personality judgment accuracy did not increase significantly over time for all big five traits. After five minutes of conversation, accuracy increased significantly for extraversion and less so for openness. Neuroticism judgments were initially accurate and stayed accurate after five minutes of conversation. However, conscientiousness and agreeableness accuracy decreased from zero acquaintance to five minute interaction.

Previous studies (Borkenau & Liebler, 1992) have reported high accuracy for extraversion at zero acquaintance and low accuracy for other traits, but this was not the case in the current study. However, the results for five minute interaction extraversion accuracy coincided with previous findings for zero acquaintance judgments. This may be due to the conditions of the experiment as participants were instructed specifically to not talk or communicate during the zero acquaintance. Extraverts are described as talkative (Costa & McCrae, 1985), so asking judges to rate extraversion in a context where one of the major cues is not available may decrease accuracy.

The fact that extraversion accuracy appeared after five minutes of interaction but was not present at zero acquaintance shows us that extraverted behavior was not completely hidden but for some reason was not detectable in the beginning of the relationship. However, it has been shown that extraversion is detectable in still photos

(Borkenau & Liebler, 1992), therefore the absence of extraversion accuracy within the current zero acquaintance paradigm using live subjects is perplexing. A possible explanation for this would stem from the uncomfortable situation the participants were subjected to for their zero acquaintance ratings, stifling any perceptible amount of outgoing behavior. A participant who would usually communicate to others in an awkward situation to relieve the tension was not allowed to do so, possibly changing the typical manifestation of their behavior and muddying the extraversion ratings by judges. The strict instructions not to communicate may have led participants to tone down extraversion cues that may normally be displayed.

These results can be interpreted using Funder's Realistic Accuracy Model (RAM) for personality judgment (Funder, 1995). Funder's RAM consists of four stages: availability, relevance, detection, and utilization. It states that in order for personality traits to be judged, the behavioral cues that reveal these judgments must be relevant to the traits being judged as well as be available to the perceiver (See Figure 2). Once cues are relevant and available, they must not only be detected by the judge, but utilized correctly. If a judge believes a cue relates to a given trait and in fact it does not, the judge did not use the cue correctly and perfect accuracy cannot be attained.

Figure 2. Realistic Accuracy Model
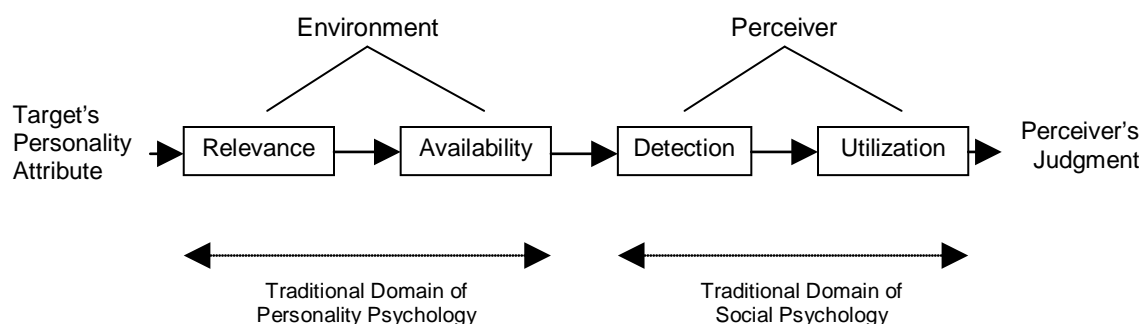
When looking at level of acquaintance with the RAM in mind, it is clear how it may affect accuracy. Merely talking with another individual for five minutes may elicit more behaviors on which to base judgments. And, given the cues are relevant, there is increased opportunity to detect and use them, integrating the cues into the perceiver's judgment.

The fact that neuroticism was detectable at zero acquaintance could have been a result of the experimental paradigm. Behavior exhibiting neuroticism may have been increased by the context of these interactions. According to Funder's (1995) RAM, the environment plays an important role in how judgments are made by a perceiver. The awkwardness of the situation (i.e. strict instructions not to communicate) while making zero-acquaintance ratings could have caused participants high in neuroticism to appear particularly neurotic, even more so than previous studies. Being instructed not to communicate with other individuals, which may have also been interpreted as not behaving extraverted, may have increased the proclivity to feel negative emotions. The outward manifestation of the feelings of negativity would cause an increase in its detectability, in turn increasing neuroticism accuracy.

Accuracy increased over time for extraversion and conscientiousness and remained consistently high for neuroticism. Openness and agreeableness accuracy, however, declined after five minutes of interaction. Even though there was an increase in information, the available cues may not have been relevant to the traits being judged or may have been misinterpreted. According to Funder's (1995) RAM, an increase in the availability of cues will cause an increase in judgment accuracy, all other things staying equal. However, if the cues are not relevant to the trait being judged, accuracy will not

increase. For openness and agreeableness, it is plausible that judges were basing ratings on cues more applicable to extraversion or conscientiousness, in turn making ratings inaccurate. This suggests that more information may not necessarily be useful for less outwardly manifesting traits such as openness and agreeableness unless behavioral cues are relevant to those traits.

Absolute Zero Acquaintance

A traditional zero acquaintance paradigm is defined as a situation in which individuals have little to no prior interaction (Albright et al., 1988). In Albright et al.'s (1988) study, before making zero acquaintance personality judgments participants were asked to silently divide themselves into groups with individuals they had not previously met. This design could potentially be problematic because the acts of looking for others to be in a group with and deciding on a place to congregate allow for the display of nonverbal behaviors, especially those related to extraversion. Simply being instructed not to talk does not prevent the exchange of nonverbal communication. At zero acquaintance, one would not expect to have even this information available. The zero acquaintance situation described in the Albright et al. (1988) study may have allowed for enough relevant expressive behavior to enable an accurate judgment of extraversion. According to the RAM, having this supplemental information available may increase accuracy. If so, then it would make it more similar to the five minute interaction paradigm. Traits expected not to be apparent until after five minutes of conversation may have leaked cues during the nonverbal exchange at zero acquaintance.

It is possible that in the present study a true zero acquaintance paradigm has been created. Participants had very little nonverbal communication prior to making zero acquaintance personality judgments. They sat in assigned seats and did not have the opportunity even to look at their group members until instructed to turn their chairs facing the others and make ratings. The explicit instructions did not allow for participants to communicate with group members about their acquaintance of another participant or decisions on where to sit. Furthermore, the instructions may have implicitly led participants to actively inhibit normally occurring expressive behavior during this "silent waiting" period. Applying this to the RAM, the information available to participants is limited to merely visual cues in the moment prior to any interaction: what the zero acquaintance situation was meant to be.

Self-Rated Trait Criteria vs. Psychological Assessment Criteria

Many methodological issues surround the accuracy criterion. A study of short and long forms of the Eysenck Personality Profiler (Petrides, Jackson, Furnham, & Levine, 2003) showed that the long forms are more comprehensive and valid. Another study by Denissen, Geenen, Selfhout, and van Aken (2008) showed that the correlation between the Ten Item Personality Inventory-Revised (TIPI-r) and Big Five Inventory (BFI) was moderate, but that the more comprehensive measures were more valid for judging traits.

For the current study, we were able to look at two different methods of measurement which both have their benefits and drawbacks. First, we were able to directly relate the self ratings of participants to ratings of judges by comparing responses on the TIPI. The advantage of this method is that the targets and judges are using the

exact same rating scale, providing for easier comparison. However, there are very few items to reliably determine each trait.

For a measure as brief as the TIPI, it is apparent to the keen participant which trait each question addresses. Unfortunately, this may have led to presentation bias in the self ratings (Dunnett et al., 1981). Dunnett et al. (1981) found that when individuals are asked to rate themselves more favorably on the Eysenck Personality Inventory (EPI), they report higher extraversion and lower neuroticism. Participants in the present study rated themselves, on average, as having higher extraversion and lower neuroticism than the others in their group rated them, which suggests that participants fell subject to self presentation bias. This bias is important to be aware of when using self ratings as an accuracy criterion.

Future Considerations

Accuracy in judging extraversion is typically present at zero acquaintance, but was not seen in this study. Extraversion showed an increase in accuracy for the five minute interaction judgments. This suggests the methodology behind the traditional zero acquaintance paradigm may not be reflecting as much "zero acquaintance" as once thought. Allowing participants to designate their own groups may have led to the exchange of nonverbal information. The situation in the current study may actually reflect "absolute zero", the time at which individuals first make eye contact with one another; the time before individuals are allowed to interact or communicate either verbally or nonverbally. Specifically, extraversion may not be as visible at true zero acquaintance as was suggested by previous research that may have unwittingly allowed

the leakage of expressive nonverbal behavior relevant to the trait extraversion. This finding alone could have major implications for the study of first impression contexts. Further research needs to be conducted to determine where the methodological issues lie.

It was also determined that the TIPI and NEO PI-R were significantly correlated for all traits except for openness, essentially replicating previous findings (Gosling et al., 2003). This suggests that in research where there is not sufficient time or resources to use a full personality inventory such as the NEO PI-R that the TIPI is a suitable replacement.

It is becoming more apparent that one of the major moderators of judgment accuracy is the degree to which two individuals are acquainted. After five minutes of conversation, a dyadic interaction is not considered a relationship. This interaction is merely a thin slice of what may turn into a relationship, the very first thin slice (Ambady, Bernieri, & Richeson, 2000). There is still a long way to go in the determination of judgment accuracy measurement but it is the hope that this study was a step in the right direction.

Making judgments of others is a reciprocal process we are involved in daily without realizing it. First impressions influence the outcomes of further interactions and can greatly influence our perceptions of others. Accuracy in these judgments will shape any and all future interactions with these individuals, making person perception an integral part of our social functioning.

References

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55,* 387-395.

Algina, J. & Olejnik, S. (2003). Sample size tables for correlation analysis with applications in partial correlation and multiple regression analysis. *Multivariate Behavioral Research, 38*, 309-323.

Ambady, N., Bernieri, F., & Richeson (2000). Toward a histology of social behavior: Judgment accuracy from thin slices of the behavioral stream. In M. P. Zanna (Ed.) *Advances in Experimental Social Psychology, 32*, 201-271.

Ambady, N., Hallahan, M., & Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology, 69,* 518-529.

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.

Blackman, M. C. & Funder, D. C. (1996). The effect of information on the accuracy and consensus of personality judgments. *Dissertation Abstracts International: Section B: The Sciences and Engineering. 57*, 4074.

Borkenau, P. & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology, 62*, 645-657.

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

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

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

Denissen, J. J. A., Geenen, R., Selfhout, M., & Van Aken, M. A. G. (2008). Single-itme big five ratings in a social network design. *European Journal of Personality, 22*, 37-54.

DePaulo, B., Kenny, D., Hoover, C., Webb, W., & Oliver, P. (1987). Accuracy of person perception: Do people know what kinds of impressions they convey? *Journal of Personality and Social Psychology, 52*, 303-315.

Dicken, C. F. (1959). Simulated patterns on the Edwards Personal Preference Schedule. *Journal of Applied Psychology, 43*, 372-378.

Dunnett, S., Koun, S., & Barber, P. J. (1981). Social desirability in the Eysenck Personality Inventory. *British Journal of Psychology, 72*, 19-26.

Funder, D. (1987). Error and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin, 101*, 75-90.

Funder, D. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652-670.

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

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62,* 88-97.

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

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.

McLarney-Vesotski, A. R., Bernieri, F., Rempala, D. (2006). Personality perception: A developmental study. *Journal of Research in Personality, 40*, 652-674.

Merrens, M. (1975). The relationship between personality inventory scores and self ratings. *The Journal of Social Psychology, 97*, 139-140.

Petrides, K. V., Jackson, C. F., Furnham, A., & Levine, S. Z. (2003). Exploring issues of personality measurement and structure through the development of a short form of the Eysenck Personality Profiler. *Journal of Personality Assessment, 81*, 271-280.

Prickett, T. J., Gada-Jain, N., & Bernieri, F. J., (2000, May). *The importance of first impressions in a job interview.* Paper presented at the meeting of the Midwestern Psychological Association, Chicago, IL.

Rosenthal, R., & Jacobson, L. F. (1968). *Pygmalion in the classroom: Teacher expectations and intellectual development.* New York: Holt.

Sharon, A. T. & Bartlett, C.J. (1969). Effect of instructional conditions in producing leniency on two types of rating scales. *Personnel Psychology, 22*, 251-263.

Shevlin, M., Walker, S., Davies, M. N. O., Banyard, P., & Lewis, C. A. (2003). Can you judge a book by its cover? Evidence of self-stranger agreement on personality at zero acquaintance. *Personality and Individual Difference, 35*, 1373-1383.

Appendix A

Sample items from the NEO PI-R (Costa & McCrae, 1992)

Neuroticism
      I am not a worrier. (R)
      I often get angry at the way people treat me.

Extraversion
      I shy away from crowds of people. (R)
      I really like most people I meet.

Openness
      Aesthetic and artistic concerns aren't very important to me. (R)
      I have a very active imagination.

Agreeableness
      I tend to be cynical and skeptical of others' intentions. (R)
      I'm not crafty or sly.

Conscientiousness
      I would rather keep my options open than plan everything in advance. (R)
      I'm known for prudence and common sense.

Appendix B

Sample items from the TIPI (Gosling, Rentfrow, Swann Jr., 2003)

Please rate the extent to which the pair of traits applies to **Individual** _ in general (i.e., yesterday, today, and tomorrow)**.**

This individual appears to be:

1.	not extraverted, not enthusiastic	1	2	3	4	5	6	7	extraverted, enthusiastic
2.	not critical, not quarrelsome	1	2	3	4	5	6	7	critical, quarrelsome (R)
3.	not dependable, not self disciplined	1	2	3	4	5	6	7	dependable, self disciplined
4.	not anxious, not easily upset	1	2	3	4	5	6	7	anxious, easily upset
5.	not open to new experiences, simple	1	2	3	4	5	6	7	open to new experiences, complex
6.	not reserved, not quiet	1	2	3	4	5	6	7	reserved, quiet (R)
7.	not sympathetic, not warm	1	2	3	4	5	6	7	sympathetic, warm
8.	organized, careful	1	2	3	4	5	6	7	disorganized, careless (R)
9.	not calm, emotionally unstable	1	2	3	4	5	6	7	calm, emotionally stable (R)
10.	not conventional	1	2	3	4	5	6	7	conventional, not creative (R)
11.	not dominant [1]	1	2	3	4	5	6	7	dominant

[1] This item does not appear in the TIPI.

Appendix C

Emotional State items from the Zero Acquaintance ratings

Emotional State right now. For each of the following scales observe **Individual _.**
Circle the number that most closely describes how this person seems at this *current moment*.

1.	nervous	0 1 2 3 4 5 6 7 8	relaxed
2.	interested	0 1 2 3 4 5 6 7 8	bored
3.	focused	0 1 2 3 4 5 6 7 8	distracted
4.	engaged	0 1 2 3 4 5 6 7 8	blank / vacant
5.	open / frank	0 1 2 3 4 5 6 7 8	calculating
6.	stressed	0 1 2 3 4 5 6 7 8	calm / collected

Appendix D

Seeing Relationships items from the Zero Acquaintance ratings

Please circle the answer to each question, even if you feel you are merely guessing.

1) Is this person currently romantically involved?   Y   N

2) Is this person currently infatuated with (currently has a crush on) anyone?   Y   N

3) Is this person currently employed?   Y   N

4) Have the parents of this person ever been divorced?   Y   N

5) Does this person have any pets?   Y   N

6) Does this person live alone?   Y   N

7) Has Individual _ lived with any **male** (Adopted/Half/Biological/Step) siblings?   Y   N

8) Has Individual _ lived with any **female** (Adopted/Half/Biological/Step) siblings? Y   N

9) In which birth order position does this person consider themselves?   Oldest   Middle   Youngest   Only

Appendix E

Predicting Relationships items from the Zero Acquaintance ratings

Please answer the following questions regarding your potential relations with the other members of your group. Some of the questions may have an intimate/personal connotation, and we ask that you please answer the questions as honestly as possible. The members of your group will not know how you rate them.

CIRCLE 1 – 8

On the last day of class...	**Not At All**							**Very Much**
1) How much will _ like you?	1	2	3	4	5	6	7	8
2) How much will you like _?	1	2	3	4	5	6	7	8
3) How interested is _ in getting to know you better?	1	2	3	4	5	6	7	8
4) How interested are you in getting to know _ better?	1	2	3	4	5	6	7	8
5) How interested will _ *say* they are in getting to know you better?	1	2	3	4	5	6	7	8
6) How intimate/close will _ *say* they are to you?	1	2	3	4	5	6	7	8
7) How sexually attracted will _ be to you?	1	2	3	4	5	6	7	8
8) How sexually attracted will you be to _?	1	2	3	4	5	6	7	8
9) How sexually attracted will _ *say* they are to you?	1	2	3	4	5	6	7	8

Footnotes

[1] Statistical significance was assessed by comparing the means from this sample with the population means reported on the personality profiles. The sample mean for male neuroticism fell within one standard deviation of reported means, indicating the males in this sample were still within the "average" range for neuroticism.

[2] The judgment accuracy coefficient was calculated for each judge by correlating personality trait judgments with traits measured for each target by the NEO PI-R. The correlation was transformed to Fisher zs for analysis.

[3] No accuracy in this context is equal to zero. A t-test against zero indicates that the level of accuracy achieved was significantly greater than chance.