Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Journal of Research in Personality 44 (2010) 38–45

Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

The effects of information and exposure on self-other agreement

Andrew Beer [a,*], David Watson [b]

[a] University of South Carolina Upstate, Psychology, 800 University Way, Spartanburg, SC 29303, USA
[b] University of Iowa, Iowa City, USA

ARTICLE INFO

Article history:
Available online 8 October 2009

Keywords:
Self-other agreement
Person perception
Zero acquaintance
Personality assessment
Information
Exposure
Acquaintanceship

ABSTRACT

The primary aims of this study were to examine the effects of exposure and information on self-other agreement among strangers. To test the effect of exposure, we observed the difference between perceptions of individuals who viewed a still photo of a target person versus individuals who watched a short video segment of the target. To test the effect of information, we provided trait-implying sentences to participants and compared the resulting perceptions to those derived from only a still photo. We found that self-other agreement fluctuates predictably with additional information and exposure. Also, we found that providing specific trait information can increase self-other agreement for both the specific trait about which information has been received and other linked trait judgments.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

The path to accurate interpersonal personality assessment is a difficult one, yet many daily interactions require an attempt to traverse it. We may rely on subtle cues presented in the physical appearance of others or we may depend on a secondary source of information, such as another's statement about the individual in question, to arrive at what we hope is the most accurate assessment of a target individual's general traits and tendencies. To date, little research has addressed similarities and differences in how information, whether direct or indirect, affects personality judgment early in the person perception process. Although we know that people who are well-acquainted show better agreement than unacquainted individuals, we still have yet to determine the specific mechanisms that give rise to this effect. The current research examines some fundamental issues regarding the role of information and observation in establishing an understanding of another person.

1.1. Factors influencing self-other agreement: acquaintanceship and visibility

How well do judges and targets agree on the assessment of the target's personality? A basic conclusion is that well-acquainted individuals typically show good agreement across major trait dimensions ($r > .40$; Funder & Dobroth, 1987; Paunonen, 1989; Watson, Hubbard, & Wiese, 2000). Of course, myriad factors influence levels of agreement. The acquaintanceship effect refers to the tendency for self-other agreement in personality judgments to increase with increased familiarity with the target. Greater acquaintanceship implies more opportunities for relevant behaviors to be exhibited by the target and attended to by the judge. This effect has been well supported in the literature (Funder & Colvin 1988; Norman & Goldberg, 1966; Paulhus & Bruce, 1992; Watson et al., 2000), though some researchers claim that whether or not there is an acquaintanceship effect depends upon whether the research design is cross-sectional or longitudinal (Kenny, Albright, Malloy, & Kashy, 1994) and to what extent the trait in question is readily observable (Paunonen, 1989).

One problem with many cross-sectional demonstrations of the acquaintanceship effect is that the groups being compared are likely to differ on variables other than acquaintanceship. For example, in Norman and Goldberg's (1966) classic study, the two contrasted groups were Peace Corps trainees and fraternity members. These groups may differ systematically in age, ethnicity, diversity, intelligence and even relevant personality characteristics. Thus, it is necessary to attempt to equate those characteristics as much as possible in order to isolate specific mechanisms of acquaintanceship.

This, then, leads to the question: What exactly constitutes "acquaintanceship"? Even if the groups are relatively equivalent aside from acquaintanceship, one still must have a formal definition of the construct. Is it simply relationship duration, or is there a qualitative difference in information gained depending upon the type of relationship? For example, is there something different about a married couple that would lead to greater agreement beyond the mere length of acquaintance? In this regard, some

* Corresponding author.
E-mail address: abeer@uscupstate.edu (A. Beer).

0092-6566/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jrp.2009.10.002

Author's personal copy

researchers have directly asked participants to estimate their level of acquaintanceship with a target and found that this estimate correlates significantly with self-other agreement (Paunonen, 1989).

One way to bypass this issue is to look at the acquaintanceship effect as based purely on information exchange. The issue then becomes: How does one control for differential information exchange?

Borkenau and Liebler (1992) conducted one of the most comprehensive examinations of sources of validity in personality judgment at zero acquaintance. In this study, the researchers collected short samples of behavior (reading a weather report) from 100 participants and then had four separate groups of judges (six per group) rate all 100 targets on the big five personality dimensions, physical attributes, and various non-verbal behaviors. One group saw only a still photo (taken from the video segment) of the target, another group heard only the audio portion of the segment, a third group saw only the video (without audio), and a final group saw the full video segment with sound. Overall, they reported that Extraversion, Agreeableness, and Conscientiousness yield greater self-other agreement among unacquainted individuals than do other traits. Furthermore, they found that as exposure to a target increases, so does self-other agreement. In some ways, one could view this study as an experimental demonstration of the acquaintanceship effect, given that judges had varying degrees of exposure to the target (e.g., those who saw the full video received more information than those who rated the still photo) and self-other agreement tended to increase with greater exposure.

Recently, research in this area has gained momentum. Carney, Colvin, and Hall (2007) controlled the quantity of information exchange by exposing participants to either 5 s, 20 s, 45 s, 60 s, or 300 s of videotaped target behavior and found a general trend towards increased self-other agreement with greater length of exposure. Letzring, Wells, and Funder (2006) investigated differences in both quantity and quality of information. They found that both short (50 min) and long (3 h) unstructured conversations lead to greater self-other agreement than a minimal contact (no verbal interaction) condition. In terms of information quality, they found that situations in which individuals were instructed to get to know each other as well as possible led to greater agreement than situations that were either unstructured or structured in such a way as to steer conversation away from personal information (a trivia quiz). Thus, previous research establishes that both quality and quantity of information have unique influences upon self-other agreement in personality perception; this study therefore aims to account for both quantity and quality of information in a slightly modified format.

2. The current study

This study examines the relative contributions of (a) increased exposure to a target and (b) increased information about a target to self-other agreement. First, replicating Borkenau and Liebler (1992), we used a still photo condition in which the participants viewed only a photo of the target and were asked to make global personality judgments. Then, in a second condition, we examined the effect of valid trait information, rather than simple exposure, on agreement by providing relevant personality trait information about targets prior to assessment. Participants were given a photo of a target, and in addition, they received a one sentence behavioral description that reflected the target individual's standing on the trait dimension of Agreeableness. This information was valid, that is, based upon actual responses from the target in a previous assessment (Beer & Watson, 2008b). A third condition specifically varied the amount of exposure a judge has to any target. Participants viewed both a photo and a short videotaped segment of the target's behavior. Thus, the video condition represents a pure difference in the quantity of information, whereas the information condition represents a difference in both quantity and quality.

Finally, in an attempt to explicate the mechanisms through which these judgments are determined, we conduct a lens model (Brunswik, 1956) analysis across the three conditions using the targets' assessed physical attributes as cues. Brunswik's model posits that there are features of a person or his or her environment that serve as a lens through which we indirectly view the core features of that person. For example, we may use our perceptions of something as basic as height to determine something about an individual's level of Neuroticism (e.g., tall people are calm). That link is referred to as cue utilization, which is the correlation between a judgment made by an observer and an objective rating of the cue. The second question of consequence concerns the issue of cue validity, or the extent to which objective ratings of the cue correlate with the actual value of the underlying construct. To continue the earlier example, cue validity is the extent to which the target person's actual level of Neuroticism is associated with that person's height. Ideally, these two correlations (validity and utilization) should match closely in terms of sign and direction. When they do, they are likely to yield an accurate judgment. If we see tall people as calm, and tall people are actually calm, then we can state that the cue of height is both valid and properly utilized. However, it may be the case that we see tall people as calm, whereas tall people actually are quite anxious. Here, the cue would still have validity, as height is related to the underlying construct, but we would have a case of improper utilization. A cue only lacks validity if it is totally unrelated to the underlying construct in question.

2.1. Hypotheses

Hypothesis 1. It was predicted that the self-other agreement correlation for Agreeableness in the information condition should exceed that in the still photo condition. Simply put, the judges are being provided valid trait information in the information condition, and they were expected to use this information to aid in their judgments. In addition, given that there is a moderate relation among peer judgments of Agreeableness, Conscientiousness, and Neuroticism (Beer & Watson, 2008a; Digman, 1997; Markon, Krueger & Watson, 2005), one would expect to see more modest gains in self-other agreement for Conscientiousness and Neuroticism in this condition.

Hypothesis 2. As established in previous work (Ambady & Rosenthal, 1992; Borkenau & Liebler, 1992), we expected that self-other agreement correlations would tend to increase as level of exposure increased from still photo to videotaped segment, demonstrating a laboratory version of the acquaintanceship effect. In periods of time as short as 30 s, individuals are able to view sufficient nonverbal behavior to increase judgmental accuracy (Ambady & Rosenthal, 1992). Previous work (Borkenau & Liebler, 1992) has demonstrated that under these conditions, agreement will increase for Extraversion, Agreeableness and Openness to Experience/Intellect, but not for Neuroticism or Conscientiousness.

3. Method

3.1. The target pool

Beer and Watson (2008b) conducted a study of stranger ratings in which each participant served as both a judge and a target in a round-robin exercise in which they assessed personality and various physical attributes. In addition, each participant was

Author's personal copy

photographed (making a "neutral" facial expression) and was asked to read a statement while being videotaped. The video segment consisted of the participant walking up to a table, sitting down, and reading verbatim a passage taken from the University of Iowa's welcome statement (full text provided in Appendix A). Each participant was given a few minutes to read the passage silently before the recording, and each recording was conducted privately (in a separate room from the rest of the experimental group). Participants were told only to read the passage, and were given no further instructions or explanation for the task. The video segments averaged approximately 45 s in length. There were 53 total groups (12 groups of 3, 23 groups of 4, and 18 groups of 5). Participants consisted of 218 (48 male and 170 female) undergraduates from an introductory psychology course at the University of Iowa and comprise the pool of available targets for the current study. Thus, for each target, we have a digital still photo and a video clip. We also have self-reports and aggregated peer reports of personality using the same Big Five scales described later. Finally, we obtained ratings on 10 physical attributes (e.g., attractiveness, height, hair color, style of dress) for each target. These attributes were rated by each member of the groups (3–5 members) using 7-point bipolar scales (e.g., 1 = light hair, 7 = dark hair). These peer ratings were then averaged to form a more objective aggregated peer evaluation of each attribute.

From this pool of 218 possible targets, five groups of 10 were selected to be used in the current study. This number was chosen specifically for two reasons: (a) it is important to examine responses across a range of targets rather than simply one or two (Wells & Windschitl, 1999) and (b) a judge could conceivably make a self-rating and approximately 10 sets of peer ratings in the span of a 1 h session. The selections were examined to ensure that the target data were consistently of high quality (no random responses, incomplete data, or low quality pictures or videos). If there was a problem along these lines, another target was selected to replace the problematic case.

Due to the manipulation of trait information, stimulus selection depended in part on the target's score on the Big Five dimension of Agreeableness. Scores on this trait dimension for each target were converted into percentile ranks from the entire initial target pool of 218 participants. For each set of 10 targets, we selected four targets scoring at or above the 70th percentile in Agreeableness, four targets scoring at or below the 30th percentile in Agreeableness, and two targets scoring between the 40th and 60th percentile in Agreeableness. If there were any problems such as those described above, a new target would be selected at random from whichever Agreeableness grouping the previous target was drawn. The final pool of 50 targets did not differ significantly from the initial pool of 218 participants in terms of self-ratings on any of the Big Five personality traits.

3.2. Participants

Newly recruited participants in this study will hereafter be referred to as judges, although we also collected their self-ratings on these same traits. The judges were selected from the research participant pool at the University of Iowa. As the study required no prior acquaintance between judges and targets, it was conducted at least 1 year after the target data were collected. In addition, judges were instructed that if at any time during the study they encountered a target with whom they were acquainted in any capacity they were to skip rating that target.

The still photo condition consisted of 154 judges (109 female), the information condition consisted of 156 judges (114 female), and the video condition consisted of 161 judges (109 female) for an overall total of 471 judges. The still photo condition will serve as the baseline reference point for the other conditions.

3.3. Measures

Each judge rated him/herself and the 10 targets on the Big Five personality traits. To assess personality, we used an abbreviated Big Five instrument designed specifically for instances in which individuals may be asked to make multiple ratings (Saucier, 1994). This 40-item instrument was developed from Goldberg's (1992) 100-item inventory and consists of adjectives that tap Emotional Stability versus Neuroticism (e.g., *moody, jealous and temperamental*), Intellect or Openness (e.g., *creative, imaginative and philosophical*), Extraversion (e.g., *talkative, bold and energetic*), Conscientiousness (e.g., *organized, efficient and systematic*), and Agreeableness (e.g., *sympathetic, warm and kind*). The participants rated themselves using a five point response format (1 = very uncharacteristic of me, 5 = very characteristic of me) for each adjective. Saucier (1994) reported that coefficient alphas for the scales ranged from .78 to .83 and from .76 to .85 for self- and liked-other ratings, respectively.

In the peer-rating version of the instrument, a blank was provided at the top of the page for the participant to designate, using the assigned number and letter, which target they were rating on that page. The instructions were modified slightly, asking the participant to "please indicate the extent to which you feel this word characterizes the individual you have indicated in the space above". The participants responded using a five point rating scale (1 = very uncharacteristic of him/her, 5 = very characteristic of him/her) for each adjective.

Coefficient alphas for the self-ratings across all conditions were .76 for Neuroticism, .86 for Extraversion, .74 for Intellect, .85 for Agreeableness, and .83 for Conscientiousness. Alphas for the peer ratings were calculated separately within each condition. Within each condition, alphas were calculated for each of the 10 peer ratings and then averaged to arrive at a single summary index of reliability. Coefficient alphas for peer ratings ranged from .77 to .92 in the still photo condition, from .75 to .95 in the information condition, and from .76 to .91 in the photo/video condition.

In addition to the personality inventory, participants were also asked to rate their anticipated liking for each target on a 1–7 scale (1 = I would not like this person very much, 7 = I would like this person very much).

3.4. General analytic strategy

This study primarily involves correlational analyses, despite the experimental manipulation. Self-other agreement was measured by determining the correlation between each individual judge's rating of each target and that target's self-rating. Thus, the number of observations in the calculation of the estimate was approximately 10 times the actual sample size. In other words, each separate target rating was treated as an individual case, rather than one-tenth of an individual case, which results in the problem of non-independent observations. For example, we had a total of 1522 observations in the still photo condition (154 judges $\times$ 10 targets per judge, minus 18 missing data points). However, we compared these estimates across conditions via a test for differences between correlations obtained from independent samples using the actual sample size (i.e., the actual number of judges) to determine the degrees of freedom, in order to eliminate potential problems related to non-independence; thus, for instance, we used a sample size of 154 for significance tests involving the still photo condition.

3.5. Procedure

Groups of judges participated in 1 h sessions. They first rated themselves on the personality questionnaire. The participants then

Author's personal copy

were instructed to use the information provided (which varied across condition) to make their best attempt at judging the personality of the targets.

The still photo condition consisted of the judges viewing photographs of the targets without any additional information, and making personality judgments based only on the photo. As in the other conditions, the photo of the target was displayed on a projection screen.

In the information condition, the judges were shown a photo of the target, together with a behavioral sentence (adapted from Uleman, 1988) that accurately reflected that individual's percentile rank on Agreeableness. For example, if the target ranked in the 80th percentile on Agreeableness in the initial Beer and Watson (2008b) sample, then the sentence reflected the fact that this individual was quite agreeable. Each target was paired with a different behavioral sentence, but all sentences reflected the dimension of Agreeableness. These sentences are included in Appendix B.

The four main stem sentences (e.g., "She is the kind of person who... likes to leave very large tips at restaurants; ...often gives food and money to beggars; ...gives a generous portion of her salary to charity; ...frequently goes out of her way to help friends") were adapted from Uleman (1988), and reflected items that correlated strongly with ratings of Agreeableness. We then created the middle and low Agreeableness sentences by adjusting the phrasing to suggest either an average or relatively low standing on the dimension. Thus, in the information condition, the four high-scoring targets were paired with sentences that would ostensibly lead to higher ratings of target Agreeableness, and these sentences were reversed for the four low-scoring targets. The two middle scorers were paired with sentences that suggested an average level of Agreeableness (e.g., She is the kind of person who leaves average tips at restaurants). After viewing the photo, the judges rated each target's personality using the same instrument they used to make their self-ratings.

In the video condition, the video of the target was played on a projection screen (lasting approximately 45 s), and then the judges were given 5–6 min to rate the personality of the viewed target. During the rating, the participants were able to look at a still photo of the target. When all judges had finished rating a target, the next video was displayed.

3.5.1. Manipulation check for the information condition

Results indicate that these sentences were an effective manipulation in that mean Agreeableness scores were significantly higher when the behavioral sentence implied high Agreeableness and lower when the behavioral sentence implied low Agreeableness. A repeated measures ANOVA for peer-rated Agreeableness revealed a main effect for Sentence Level, $F(2310) = 506.30$, $p < .0001$; follow up contrasts revealed that low Agreeableness sentences ($M = 20.82$, SD = 4.07) yielded significantly lower Agreeableness ratings than did the middle Agreeableness sentences ($M = 30.98$, SD = 4.82), $F(1155) = 424.50$, $p < .0001$, and that middle Agreeableness sentences yielded significantly lower Agreeableness ratings than did high Agreeableness sentences ($M = 34.40$, SD = 3.21), $F(1155) = 89.22$, $p < .0001$.

4. Results

4.1. Descriptive statistics

Before turning to our formal hypotheses, we briefly examine the descriptive statistics and intercorrelations among the trait ratings to help clarify some of the results reported later. Table 1 provides the means and standard deviations for the Big Five ratings in each condition. There was a general trend towards more favorable self-ratings than peer ratings across all three conditions. Effects were strongest for Agreeableness in every condition ($d = 1.47$ in the still photo condition, 1.60 in the information condition, and 1.35 in the video condition).

Table 1
Descriptive statistics.

Variable	Self		Peer		t	d
	Mean	SD	Mean	SD		
Still photo ($N = 154$)						
Neuroticism	20.61	5.00	22.36	2.74	−4.20**	−.43
Extraversion	29.21	5.70	24.36	2.59	9.09**	1.10
Intellect	27.81	4.80	25.36	2.42	5.93**	.64
Agreeableness	33.97	5.20	27.60	3.27	15.45**	1.47
Conscientiousness	30.32	5.73	27.29	2.70	5.91**	.68
Information ($N = 156$)						
Neuroticism	20.13	4.80	21.81	2.29	−4.53**	−.45
Extraversion	29.08	5.73	26.31	2.45	5.82**	.63
Intellect	28.58	4.38	25.56	2.21	8.56**	.87
Agreeableness	34.02	4.50	28.29	2.32	16.51**	1.60
Conscientiousness	30.82	5.06	27.74	2.39	7.45**	.78
Video ($N = 161$)						
Neuroticism	20.30	5.16	21.02	2.86	−1.59	−.17
Extraversion	28.79	6.17	25.05	2.36	7.02**	.80
Intellect	28.18	4.66	24.82	2.54	8.12**	.90
Agreeableness	33.93	4.46	28.85	2.91	13.40**	1.35
Conscientiousness	31.00	4.88	28.09	2.80	6.70**	.73

** t is significant at $p < .01$, two-tailed.

4.2. Intercorrelations among the trait ratings

Table 2 provides the intercorrelations among the self and peer ratings. The latter are collapsed across all conditions due to a lack of any significant differences between them. Consistent with previous research (see Beer & Watson, 2008a), peer judgments are less independent than self-judgments. In fact, all of the self versus peer correlations differed significantly from one another in the overall sample, with two exceptions: Extraversion–Neuroticism and Neuroticism–Intellect. These elevated correlations in the peer ratings have important implications for our information manipulation, in that the receipt of a single piece of information may, in fact, be broadly used and lead to substantial changes in linked trait judgments.

4.3. Self-other agreement

Table 3 provides the self-other agreement correlations in each condition. In the still photo condition, only one agreement correlation (Extraversion, $r = .30$, $p < .05$) is significant. This replicates Borkenau and Liebler's (1992) finding and highlights the fact that significant self-other agreement can be obtained on this highly

Table 2
Intercorrelations among self versus peer judgments.

Scale	N	E	I	A	C
Neuroticism		−.27	−.10	−.63**	−.49**
Extraversion	−.22		.25**	.36**	.27**
Intellect	−.06	.03		.33**	.47**
Agreeableness	−.32	.20	.10		.65**
Conscientiousness	−.25	.04	−.01	.40	

Note $N = 470$: self-rating correlations are below the diagonal, peer-rating correlations are above the diagonal.
** The difference between the self and peer correlations is significant at $p < .01$, two-tailed.

Author's personal copy

Table 3
Self-other agreement correlations.

	Still photo (N = 154)	Information (N = 156)	Video (N = 161)
Neuroticism	.13	.25**	.09
Extraversion	.30**	.32**	.42**
Intellect	.00	.05	−.01
Agreeableness	.09	.68**	.26**
Conscientiousness	.12	.21**	.07

Note: See methods section for details regarding the calculation of significance tests.
** $p < .01$, two-tailed.

visible trait even when limiting information to a single still photograph. In contrast, in the information condition, four of the five agreement correlations are significant (with Intellect being the exception). In accordance with Hypothesis 1, self-other agreement for Agreeableness was significantly elevated in the information condition versus the still photo condition (.68 versus .09, $z = 6.45$, $p < .01$, one-tailed). It was also predicted that we might observe more modest elevations in traits related to Agreeableness, specifically Neuroticism and Conscientiousness. Although the increases in agreement for these traits were in the predicted direction—greater self-other agreement in the information versus the still photo condition—neither difference reached statistical significance (Neuroticism: .25 versus .13, $z = 1.09$, n.s.; Conscientiousness: .21 versus .12, $z = .81$, n.s.).

In the case of the information manipulation, it appears that participants properly used the Agreeableness information provided to judge Agreeableness more accurately. It is noteworthy that four of the five agreement correlations are significant, compared to only one significant agreement correlation in the still photo condition. As expected, direct trait information does lead to gains in agreement both on the target trait and on related trait judgments.

Most notable in the video condition are the significant agreement correlations for Agreeableness and Extraversion. These results replicate Borkenau and Liebler's (1992) findings regarding the value of even limited behavioral observation for the judgment of these two traits. Hypothesis 2 predicts that self-other agreement correlations will increase for Extraversion, Agreeableness, and Intellect, and will not be affected for Neuroticism and Conscientiousness. These predictions were partially supported. Although neither of the differences reached statistical significance, both Extraversion and Agreeableness showed modest increases in self-other agreement in the video condition as compared with the still photo condition (Extraversion: .42 versus .30, $z = 1.21$, n.s.; Agreeableness: .26 versus .09, $z = 1.54$, n.s.). Intellect (−.01 versus .00, $z = -.09$, n.s.), Neuroticism (.09 versus .13, $z = -.36$, n.s.) and Conscientiousness (.07 versus .12, $z = -.44$, n.s.) all showed relatively little change across conditions. Overall, these data suggest that judges gain some valid information relevant to Extraversion and Agreeableness simply from observing a target reading a short passage.

4.4. Cue utilization

Finally, we report analyses based on Brunswik's (1956) lens model. Table 4 presents the cue validity and utilization correlations, by condition, for the five factors using 10 peer-rated attributes (see our earlier discussion of the lens model for definitions of cue validity and cue utilization). The most valid cues, overall, were fashionable dress (four significant correlations with self-rated Big Five traits), physical attractiveness (3) and stylish hair (2). As would be anticipated from the self-other agreement correlations, Extraversion was the trait that showed the most relations with observable cues, showing significant correlations with 4 of the 10 measured attributes. Only Intellect failed to show a significant validity correlation.

Table 4
Cue validity and utilization.

Cue validity					Attribute	Cue Utilization															
						Still photo					Information				Video						
N	E	I	A	C		N	E	I	A	C	N	E	I	A	C	N	E	I	A	C
−.15	.01	−.01	.09	.06	Formal dress (.67)	−.09	.11	.09	.07	.12	−.11	.09	.08	.09	.14	−.08	.05	.05	.06	.07
−.32	.37	.05	.33	.31	Fashionable Dress (.50)	−.07	.23	−.02	.02	.01	−.22	.24	.09	.21	.08	−.07	.25	−.02	.05	−.04
.12	.01	.02	−.14	−.01	Dark hair (.92)	−.02	−.06	.02	−.05	−.08	.00	−.08	−.01	−.07	−.04	−.04	−.01	.03	−.04	−.04
−.32	.24	.12	.26	.29	Stylish hair (.20)	.00	.25	−.02	−.03	−.02	−.16	.25	.05	.17	.00	−.03	.20	−.04	.02	−.07
−.12	.07	.00	−.04	.05	Tall (.75)	−.20	.08	.04	.15	.05	−.02	−.02	−.06	−.05	−.05	−.15	.26	.09	.10	.04
−.08	.33	.14	.00	.06	Muscular (.61)	−.08	.18	−.12	−.06	−.18	−.10	.13	−.08	−.03	−.18	−.10	.24	−.07	−.04	−.12
.00	−.25	−.06	.09	.00	Elongated face (.56)	−.12	−.03	.10	.11	.11	−.10	−.05	.07	.16	.11	−.11	.08	.07	.11	.10
.20	−.03	.06	−.17	−.17	Older-looking face (.13)	−.09	.11	.10	.10	.07	.05	−.02	.01	−.09	−.04	.00	−.05	−.05	−.08	−.11
−.19	.38	.23	.09	.16	Athleticism (.42)	−.18	.25	−.09	.09	−.09	−.23	.21	−.03	.11	−.09	−.15	.34	−.07	.03	−.08
−.42	.35	.02	.40	.25	Attractiveness (.51)	−.02	.29	.01	.18	.19	−.17	.29	.12	.28	.16	.05	.15	.00	.11	.07
M																				
.20	.21	.07	.16	.14		.09	.16	.06	.09	.09	.12	.14	.06	.13	.09	.08	.16	.05	.06	.07

Note: Cue validity is the correlation between aggregated peer ratings from Beer and Watson (2008b) on the physical cues and 50 targets' self-ratings on the Big Five. Cue utilization is the correlation between aggregated peer ratings from Beer and Watson (2008b) on the physical cues and the peer judgments made by judges in the current sample, calculated separately for each condition. Underlined correlations are significant at $p < .05$. The bottom row contains the mean absolute value of the columns above. Coefficient alphas for the cues rated by at least four peers ($N = 88$) are provided in parentheses.

Author's personal copy

Table 5
Column-vector correlations between cue validity and cue utilization correlations.

Scale	Condition		
	Still	Info	Video
Neuroticism	−.21	.82	−.15
Extraversion	.87	.85	.71
Intellect	−.79	−.51	−.83
Agreeableness	.19	.96	.60
Conscientiousness	.02	.30	.08
Overall	.44	.76	.46

Note: Each of the individual correlations across the conditions is based on $N = 10$ (the number of cues assessed). The "overall" category collapses across traits within each condition, computing a hetero-trait aggregated utilization correlation based on $N = 50$ (10 cues per each of the five traits). In all cases, positive correlations indicate effective cue utilization, zero correlations indicate ineffective cue utilization, and negative correlations indicate grossly ineffective cue utilization.

There certainly are some valid cues available to judges, but the question is: Do the judges utilize these cues (a) at all and (b) properly? The answer to the first question appears to be yes, with some qualifications. Each condition yielded several significant utilization correlations–between 5 and 9 across the three conditions, with the video condition showing the least utilization of the measured attributes. Across all conditions, Extraversion was the trait most often significantly associated with these cues. However, this does not necessarily mean that judges actually considered these cues prior to making a judgment. Thus, the answer to the first question: "Do judges utilize cues at all?" is not fully addressed by these correlations.

Regardless, the second question is still interesting: Was the utilization effective? Moreover, was it differentially effective across conditions? To measure this, we computed column-vector correlations for each condition (see Borkenau & Liebler, 1992; Funder & Sneed, 1993; Gosling, Ko, Mannarelli, & Morris, 2002). To do this, we first transformed the cue validity and cue utilization correlations from Table 4 using Fisher's *r*-to-*z* formula. After this, we correlated these transformed correlations by trait, across the cues. These values are reported in Table 5. For example, the −.21 utilization coefficient for Neuroticism in the still photo condition represents the correlation between the values in the first column of the left side of Table 4 and the first column of the right side of Table 4 under the heading "Still photo". Note that a near-zero or negative utilization correlation indicates improper utilization by the judges, whereas a positive utilization correlation indicates a correspondence between cue validity and cue utilization.

The most notable aspect of the data in Table 5 is that the information condition yields the most effective cue utilization. This means that in this condition in particular, judges seem to be utilizing the valid cues properly while ignoring the invalid cues. The one major exception is Intellect, a trait that shows poor cue utilization across all three conditions. It should be noted, however, that the cue validities are uniformly low, thus precluding highly effective utilization. Of greater interest is the general increase in good use of cues relevant to Neuroticism and Agreeableness in the Information condition. Viewed in conjunction with Table 4, it seems that when participants receive a piece of trait relevant information, it serves to depress usage of invalid cues and increase usage of valid ones.

5. Discussion

Personality judgments under conditions of limited acquaintance typically are largely inaccurate (Beer & Watson, 2008b; Watson, 1989). However, the literature suggests that Extraversion can be judged effectively even in such conditions (for a recent example, see Borkenau, Brecke, Mottig, & Paelecke, 2009). Based on these previous findings, we expected that Extraversion would show significant self-other agreement at the lowest levels of acquaintanceship. We aimed to determine what effect, if any, different types of information might have on self-other agreement at zero acquaintance. For the information condition, our expectation was that valid information would lead to gains in agreement, specifically for the trait relevant to the provided information (Agreeableness). We also expected more moderate gains in other trait judgments due to significant correlations among the Big Five traits, particularly within strangers' ratings (Beer & Watson, 2008a). In the video condition, we expected to observe some increases in self-other agreement, particularly for more visible traits such as Extraversion (Borkenau & Liebler, 1992), but not for less visible traits such as Neuroticism. Quite simply, increased exposure to the target should provide the judge with more valid trait information for those traits that can be expressed in such circumstances.

The still photo condition simply provided judges with a visual representation of the target. Nevertheless, we obtained an agreement correlation of .30 for Extraversion, replicating Borkenau and Liebler's (1992) finding. None of the other Big Five traits showed any agreement in this condition (*r*s ranged from .00 to .13). Together with findings from previous studies, these results again demonstrate that Extraversion is a remarkably easy trait to judge in another.

When judges received a sentence accurately implying a target's standing on the dimension of Agreeableness, we observed a huge leap in self-other agreement relative to the still photo condition (.68 versus .09) for this trait, as one would expect. In addition, we observed gains in agreement and significant self-other agreement correlations for every trait save Intellect. It is almost as if receiving this one valid statement served to calibrate the entire judgment process. Although we cannot be certain as to the mechanism involved, one possibility is that judges have a fairly accurate stereotype concerning the interrelations among traits (e.g., that agreeable people tend not to be highly neurotic). Thus, receiving valid information about any specific trait necessarily should increase self-other agreement on other related trait dimensions.

In this regard, it should be noted that although the self and peer judgments show different degrees of interrelatedness (see Table 2), the relative magnitudes of the relations among traits are fairly consistent from self to peer across conditions. In other words, those traits that show the strongest relations in self-ratings also tend to show the strongest relations in peer judgments (e.g., the correlation between Agreeableness and Neuroticism is the strongest in both the self and peer judgments). We examined the pattern of intercorrelations by condition, and found that if a column-vector correlation is computed (using a Fisher *r*-to-*z* transformation and taking the absolute values of the 10 respective self and peer correlations), the resulting coefficients are quite similar across conditions. The correlations in the still photo, information, and video conditions were .65, .76 and .65, respectively. The fact that this correlation is slightly elevated in the information condition provides further support for the notion that the supplied information related to target agreeableness helps to fine tune the perceptual process.

When judges received increased exposure to the target via a videotaped segment, we also observed increases in self-other agreement. The agreement correlation for Extraversion increased from .30 in the still photo condition to .42 in the video condition. It is noteworthy that the latter coefficient is not significantly weaker than the correlation one would observe in a friendship dyad ($r = .48$) or dating couple ($r = .45$, see Watson et al., 2000). In addition, we observed a significant self-other agreement correlation for

Author's personal copy

Agreeableness in the video condition, which again replicates results reported by Borkenau and Liebler (1992).

Overall, our predictions regarding self-other agreement were largely confirmed. An informational account of personality perception holds when using agreement as a criterion. When valid incremental information is provided to the judge, regardless of modality, we see gains in agreement.

5.1. Cue utilization

The most surprising finding with respect to cue utilization occurred in the information condition. It seems reasonable to expect that receiving a piece of valid trait relevant information would suppress a judge's utilization of all other cues, with the judge choosing to focus instead on the information in the sentence. However, we did not observe such a pattern in our data. Rather, the information provided seemed to sharpen the focus of the lens the judges were using. Having said this, however, we also should emphasize that these data are open to alternative explanations; for instance, it may be that receiving a piece of valid trait information activates stereotypes about trait intercorrelations that have some basis in fact (for an extended discussion of this point, see Kenny, 2004). More fundamentally, our design does not allow us to determine the cues the judges actually used in making their ratings.

There were fewer significant utilization correlations when participants viewed videotaped behavior than when they viewed a still photograph or a still photograph accompanied by an informative sentence. However, of the five significant utilization correlations in the video condition, three had significant corresponding cue validity correlations. So, the gains in effective utilization associated with increased exposure may be more involved with ignoring useless cues, whereas the gains associated with the receipt of trait information may be the result of both ignoring invalid cues and using the valid ones more efficiently. In addition, it seems likely that audio and dynamic visual cues unaccounted for in the current study were being more fully utilized in the video condition to the exclusion of some of the static visual cues relied upon when viewing only a still photograph. Once again, we conclude by emphasizing that we cannot know what cues the judges actually used in making their ratings.

5.2. Limitations

First, our sample consisted of undergraduates, and it was predominantly female. It is possible that there are important gender differences in personality perception that we would be unable to examine in the current study due to a paucity of male participants, both as judges and as targets. Additionally, there may be important changes in personality perception as people age that mirror general trends in personality development (Caspi, Roberts, & Shiner, 2005). In short, the nature of our sample could limit the ecological validity of the findings.

Also germane to ecological validity is the experimental setting itself. People rarely make judgments based on photographs and behavioral sentences, nor do they often see videos of individuals engaging in highly mundane behaviors such as reading a university's welcoming statement aloud. However, we would argue that these laboratory conditions are fair approximations of events that do happen in daily life. For example, one may have overheard others mentioning that a co-worker is a lousy tipper.

In addition, we would have liked to include a more extensive battery of judgments about the targets. It would be interesting to examine how information and exposure influence judgments of affectivity, attitudes, and other attributes beyond the Big Five.

Finally, our choice of Agreeableness as the trait about which judges would receive information could be criticized. Most importantly, Agreeableness is a highly evaluative trait. Thus, our manipulation became more of an instruction booklet on whom to like rather than the simple receipt of trait relevant information. We probably could have avoided this had we chosen to use either Intellect or Conscientiousness as the informative trait. However, any specific trait would have come with its own unique set of limitations; thus, the simple solution would be to replicate the study using different types of manipulated information in an attempt to tease apart potential explanations for the effects.

5.3. Future directions

The results of this study suggest several avenues of research that could prove fruitful. First, more work is necessary to determine whether different types of people perceive strangers in different ways. Much zero acquaintance perception work has been conducted using student samples, so it would be useful to examine similar phenomena in more mature adult samples.

Second, our video condition involved a very limited range of behavior (i.e., reading a mundane statement) that does not allow for much expression of the target's personality traits. It will be important for future studies in this area to follow the lead of Letzring et al. (2006) and Carney et al. (2007) and incorporate more powerful manipulations of information that provide the judge with a broader array of trait-relevant behavior.

More importantly, making judgments of personality is one thing, but the resulting decisions based on these judgments are the ultimate concern for researchers in this area. More work is necessary to clarify the link between personality judgments and social decision-making. Researchers must demonstrate that these personality judgments have real-world implications for behavior, such as hiring decisions and mate selection, in order to justify serious interest in this basic research. Such work is already underway (Olivola & Todorov, 2007), but more would certainly be welcomed.

Work in personality perception under conditions of limited acquaintance was dormant for an extended period of time, but it is gaining steam within psychology. We still understand very little of the processes underlying personality judgment, both initially and over the course of various types of relationships. Examining this phenomenon in its most basic form, stranger perception, will help us to build meaningful explanatory models with practical implications for everyday life.

Appendix A

Welcome to the University of Iowa! We are proud to be an institution whose top priorities are high-quality education, world-class research, and deeply committed service. For over 150 years, we have been educating students of all ages who have made positive impacts in their communities—within Iowa and throughout the world. Preparing our students for lives of cultural richness and good citizenship is also important to our educational mission. We have over 100 areas of study, all grounded in the liberal arts. Our graduate and professional programs offer cutting-edge training in over 100 areas, such as medicine, dentistry, law, education, business, engineering, the social sciences, the physical and biological sciences, and the arts and humanities. We were the first University to accept creative work for advanced degrees, and our Writers' Workshop remains the premier creative writing program in the world. Students here are instructed by some of the finest teachers and scholars in the world.

Author's personal copy

Appendix B

B.1. High agreeableness

She/He is the kind of person who likes to leave very large tips at restaurants.
She/He is the kind of person who often gives food and money to beggars.
She/He is the kind of person who gives a generous portion of her/his salary to charity.
She/He is the kind of person who frequently goes out of her/his way to help friends.

B.2. Middle agreeableness

She/He is the kind of person who leaves average tips at restaurants.
She/He is the kind of person who sometimes goes out of her/his way to help friends.
She/He is the kind of person who gives a moderate portion of her/his salary to charity.
She/He is the kind of person who occasionally would give food and money to beggars.

B.3. Low agreeableness

She/He is the kind of person who leaves very small tips at restaurants.
She/He is the kind of person who is unwilling to give food or money to beggars.
She/He is the kind of person who gives only a very small portion of her/his salary to charity.
She/He is the kind of person who rarely goes out of her/his way to help friends.

References

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

Beer, A., & Watson, D. (2008a). Assymmetry in personality perception: Others are less differentiated than the self. *Journal of Personality, 76*, 535–560.

Beer, A., & Watson, D. (2008b). Personality judgment at zero acquaintance: Accuracy, assumed similarity, and implicit simplicity. *Journal of Personality Assessment, 90*, 250–260.

Borkenau, P., Brecke, S., Mottig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality, 43*, 703–706.

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

Brunswik, E. (1956). *Perception and the representative design of experiments*. Berkley: University of California Press.

Carney, D. R., Colvin, R. C., & Hall, J. A. (2007). A thin-slice perspective on the accuracy of first impressions. *Journal of Research in Personality, 41*, 1054–1072.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology, 56*, 453–484.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246–1256.

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., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64*, 479–490.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26–42.

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.

Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review, 8*, 265–280.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin, 116*, 245–258.

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quality and quantity affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology, 91*, 11–123.

Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative heirarchical approach. *Journal of Personality and Social Psychology, 88*, 139–157.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees and randomness in personality structure. *Journal of Personality and Social Psychology, 4*, 681–691.

Olivola, C. Y., & Todorov, A. T. (2007). A picture is worth a thousand inferences: First impression and mate selection in internet and speed dating. In *Presentation at the 8th annual meeting of the Society of Personality and Social Psychology*, Memphis, TN.

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

Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology, 56*, 823–833.

Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar big-five markers. *Journal of Personality Assessment, 63*, 506–516.

Uleman, J. S. (1988). Trait and gist inference norms for over 300 potential trait-implying sentences. Unpublished raw data.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57*, 120–128.

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.

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25*, 1115–1125.