

Accuracy of Judging Affect and Accuracy of Judging Personality: How and When Are They Related?

Judith A. Hall,¹ Sarah D. Gunnery,² Tera D. Letzring,³
Dana R. Carney,⁴ and C. Randall Colvin¹

¹Northeastern University

²Tufts University

³Idaho State University

⁴University of California, Berkeley

Abstract

Objectives: The present research is concerned with the relation between accuracy in judging targets' affective states and accuracy in judging the same targets' personality traits. In two studies, we test the link between these two types of accuracy with the prediction that accuracy of judging traits and of judging states will be associated when fundamental affective qualities are shared.

Method: In Study 1, affective states and personality traits of 29 targets were rated by 124 judges whose individual accuracy was scored as the correlation between their ratings and target criterion scores (across targets). In Study 2, a comparable analysis was done using 30 different targets and 330 different judges.

Results: Accuracy in judging distressed affect was significantly positively correlated with accuracy in judging Neuroticism in both studies, as well as in a meta-analysis across the two studies. Accuracy in judging positive affect was significantly positively correlated with accuracy in judging Extraversion in one of the two studies, with the meta-analysis across the two studies being significant.

Conclusions: These findings provide preliminary evidence for a new model (State and Trait Accuracy Model) that outlines when concordance in accuracy across traits and states should be expected.

Keywords: interpersonal accuracy, emotion recognition, personality judgment, State and Trait Accuracy Model, self-other agreement

Within the study of accuracy of interpersonal perception, there are two traditions that have had almost no connection to each other, either empirically or theoretically. One of these traditions is mostly pursued by social psychologists who study the accurate judgment of affect or emotion, and the other is mostly pursued by personality psychologists who study the accurate judgment of personality traits. Although Bruner and Tagiuri (1954) commented long ago on the conceptual overlap between these areas, researchers still have not explored how these two skills are related overall or how specific affect judgment skills could correspond with specific personality judgment skills.

We located only two published studies that connected these two kinds of judgment skill (Ambady, Hallahan, & Rosenthal, 1995; Realo et al., 2003). In the Realo et al. (2003) study, tests of judging affect cues in face or voice did not correlate with a test of judging personality cues in a different set of targets. In Ambady et al. (1995), personality judgment accuracy was based on participants' judgments of live interactants in a zero-acquaintance situation, whereas accuracy of judging non-

personality cues was measured with two videotaped multichannel tests of decoding cues that involved affect but were not exclusively about affect or emotion recognition. One of the standardized tests was significantly positively correlated with accuracy in judging personality, but the other was not.

In the present research, we also asked about cross-domain accuracy—specifically, the correlation between accuracy of judging affect and accuracy of judging personality—but we measured accuracy when judges based both their personality and affect inferences on the *very same* social stimuli, unlike in Realo et al. (2003) and Ambady et al. (1995), where the two kinds of accuracy were measured using different test stimuli.

The authors thank Ross Corlies, PoShan Yeung, Amanda Pope, Matt Poon, Arjun Jindal, and Kailey Face for helping to conduct Study 1. Acknowledgments regarding Study 2 appear in Carney, Colvin, and Hall (2007).

Correspondence concerning this article should be addressed to Judith A. Hall, Department of Psychology, Northeastern University, Boston, MA 02115. Email: j.hall@northeastern.edu.

Basing the affect and personality judgments on the same stimuli allows for a more precise examination of the judgment process. Also, along with asking the general question of how accuracies for judging states and traits are related overall, we posed the more focused question of whether accuracy for judging *specific* kinds of affect or emotions is connected with accuracy for judging *specific* traits. Although there is almost no published research on the relation between accuracy of judging affect and accuracy for judging personality, as noted above, we could draw on theoretical logic as well as research specifically focusing on the traits of Neuroticism and Extraversion to frame both general and specific predictions.

Overall Accuracy

First, we consider the overall relation between affect judgment accuracy and personality judgment accuracy. One could predict a positive relation for two different reasons. The first would be based on the possible existence of a general factor of accuracy skill. According to this notion, if one is a good judge of other people, this should be evident for all kinds of interpersonal accuracy. Accuracy on different kinds of content to be judged (e.g., discrete emotions, personality, deception, or personal attributes of targets such as sexual orientation or nationality) should tap equally into the general factor and be positively correlated with each other because they draw on similar judgment processes and knowledge (i.e., a common skill). For many decades, scientists have posed the question of whether there is a general skill of interpersonal accuracy (e.g., Boone & Schlegel, 2016; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979; Schlegel, Grandjean, & Scherer, 2012; Vernon, 1933). A meta-analysis of correlations between tests of different kinds of interpersonal accuracy suggests that evidence for a general skill is weak, although on average the between-test correlations are positive (Boone & Schlegel, 2016).

The second, and more specific, reason for predicting a positive correlation between accuracy of judging affect and accuracy of judging personality is that accuracy in recognizing the affective states of the targets may *contribute* to accuracy in judging their personality. In other words, there could be a causal relation between these two kinds of accuracy. Past research shows that one way personality is manifest in behavior is through affective expression, both verbal and nonverbal (e.g., Augustine & Larsen, 2015; Borkenau & Liebler, 1995). Furthermore, judges infer personality from nonverbal cues, including emotional expressions (Hall, Gunnery, & Andrzejewski, 2011; Hareli & Hess, 2010; Hareli, Shomrat, & Hess, 2009; Knutson, 1996; McAleer, Todorov, & Belin, 2014; Montepare & Dobish, 2003; Petrican, Todorov, & Grady, 2014). It is logical, therefore, to posit that these skills are related because there is a causal link between accurately judging affect and accurately judging traits.

Although a general theoretical case can be made for why overall accuracy in judging affect should be related to overall accuracy in judging personality, the more precise and compel-

ling hypothesis is that accuracy of judging *specific* emotions or kinds of affect is related to accuracy of judging *specific* personality traits for which the given kind of emotion or affect is relevant. For example, a judge who believes that neurotic people are generally dysphoric (which they are; Karsten et al., 2012; Spinhoven, Penelo, de Rooij, Penninx, & Ormel, 2014) might especially rely on cues that suggest anxiety, fear, or distress in making a decision about Neuroticism. However, it should also be the case that a successful judgment of Neuroticism requires *accurate* detection and interpretation of the relevant affect cues. If accuracy of judging the affect cues is high, then this should improve accuracy in judging Neuroticism. In the present research, we give special attention to accuracy of judging two such affect–trait pairs: distressed affect and Neuroticism, and positive affect and Extraversion.

Specific Accuracies

The hypothesis we pursue in the present research is that there are *content-specific* factors that link different accuracy skills. We measured accuracy in identifying different affective states and also accuracy in identifying the Big Five traits (Extraversion, Openness to Experience, Agreeableness, Conscientiousness, and Neuroticism) to see whether there are positive correlations between accurately judging distressed affect and Neuroticism, and between accurately judging positive affect and Extraversion, based on the fact that the link between affective experience and personality is most often examined and found in relation to these two traits (Augustine & Larsen, 2015). Although finding such a correlation does not demonstrate a causal path, it would be a first step in support of a model in which accurate detection of affect influences accurate judgment of personality. Discriminant validity evidence for such a model would result if the hypothesized connections were stronger than connections involving affective states and traits for which no hypothesized connection is made (e.g., accuracy for judging positive affect correlated with accuracy for judging Conscientiousness).

Distressed Affect and Neuroticism

A salient feature of Neuroticism is the experience of fear or anxiety. Neuroticism scales emphasize distressed affect (e.g., Costa & McCrae, 1992; Gosling, Rentfrow, & Swann, 2003), and empirical research confirms this association in neurotic people (Karsten et al., 2012; Spinhoven et al., 2014). Furthermore, Neuroticism is related to nonverbal qualities of appearance and behavior; Borkenau and Liebler (1995) found that more neurotic targets behaved less positively overall (e.g., less smiling) and behaved in an uncertain and nondominant manner (e.g., less self-assured expression, less eye contact, less powerful voice, and more speech dysfluencies). Other studies have found that Neuroticism is linked to behavioral signs of anxiety or

nervousness (Back, Schmukle, & Egloff, 2009; Gawda, 2007; Hirschmüller, Egloff, Schmukle, Nestler, & Back, 2015).

It is also known that judges utilize cues of the emotional states of anxiety and nervousness to infer Neuroticism (Funder & Sneed, 1993; Hirschmüller et al., 2015), and Hall et al. (2011) found that when watching a fear-inducing film or posing fear through nonverbal cues, targets were subsequently rated by naïve judges as more neurotic than when the same targets were watching or posing different discrete emotions, controlling for the targets' actual Neuroticism levels. That study showed clearly that expressive cues of fear produce attributions of Neuroticism, even independent of whether the target is actually neurotic. All of the foregoing research lays a foundation for the hypothesis that correct interpretation and use of fear cues (or more distressed cues in general) plays a role in accuracy of judging Neuroticism.

Positive Affect and Extraversion

Research finds that positive mood states are associated with Extraversion, both in theory and according to self-report data (Augustine & Larsen, 2015; Lucas, Le, & Dyrenforth, 2008), and that more extraverted people convey nonverbal cues that are suggestive of cheerful affect (Borkenau & Liebler, 1995). However, there is also meta-analytic evidence pointing to extraverts' *general* expressivity, not just positive expressivity (Riggio & Riggio, 2002). Nevertheless, it could be reasonable to think that if the judge's implicit (or explicit) theory of Extraversion includes happy affect, then accuracy of judging happiness could be related to accuracy of judging Extraversion.

However, it is important to note that Extraversion is not conceived, either in the literature or in the popular imagination, to be entirely or even mostly about positive affect. Extraversion is above all associated with concepts like sociability, outgoingness, willingness to be the center of attention, and assertiveness especially within a social setting. For this reason, a judge's implicit (or explicit) theory of Extraversion may focus more on these interpersonal behavioral qualities than on positive affect *per se*. Therefore, although it is justified to hypothesize a relation between accuracy of judging positive affect and accuracy of judging Extraversion (Clark & Watson, 2008), it cannot be stated with as much confidence as the analogous hypothesis for the fear/distress construct and Neuroticism.

STUDY I

Method

Targets. Twenty-nine female Northeastern University undergraduate students served as targets, who participated for partial fulfillment of their course requirement in introductory psychology. The use of female targets maximizes the presence of relevant cues to affect and personality, owing to females' greater behavioral expressiveness (e.g., Hall & Friedman, 1999). The 29 targets participated with a partner who was recruited in the

same fashion and whom they did not know before coming into the lab. In order to create tapes that were short enough for judges to watch in one experimental session, one partner from each dyad was selected as a target. This partner was always the one who chose to sit in a specific chair so there was a consistent backdrop present in the background of all stimuli.

Judges. One hundred twenty-four (56% female) Northeastern University undergraduate students participated as judges of the targets' behavior; they were also recruited from introductory psychology classes for partial fulfillment of their course requirement. Judges participated by themselves or in a group of up to five people.

Target Affective Narration Task. Dyad members took turns telling their assigned partners about the time they felt the most happy, most disgusted, most fearful, and most sad. Stories were counterbalanced for order so that the dyad members did not narrate the same kind of story back-to-back. Targets were asked to narrate the happy, disgusted, fearful, and sad stories for 2 min, and the resulting stories had a mean length of 64.04 s. Both targets were videotaped during each narrative by video cameras that were placed over their partner's shoulder. This way, each target was videotaped head-on, from the waist up. (Recall, however, that only one target from each dyad was used for the judgment task.)

Target Personality Measurement. Targets' personalities were measured using the 60-item NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992), a standard measure of the Big Five personality traits (Extraversion, Conscientiousness, Agreeableness, Neuroticism, and Openness to Experience).

Target Procedure. Targets entered the lab in groups of two. First, they each completed the NEO-FFI. The experimenter then explained that the dyad would be taking turns telling about the time that they each felt the most happy, disgusted, fearful, and sad. The experimenter placed a piece of paper with the randomized order in which the targets would be telling their stories on the table in front of them. The video cameras were started, and the targets took turns telling their affect narratives. Following each narrative, the narrator filled out the affect manipulation check before listening to her partner tell her narrative. This process was repeated until both partners had completed all four narrations.

Affect Manipulation Check. Following each narration, targets self-reported how much they had felt each of 16 states on a scale of 1 (*not at all*) to 9 (*extremely*) during the preceding task. These represented happiness (*happy, content, cheerful, pleasant*), disgust (*disgusted, grossed out, sickened, repulsed*), fear (*fearful, nervous, anxious, worried*), and sadness (*sad, gloomy, depressed, downhearted*). To ascertain whether targets' reported affective state was congruent with the assigned affect of the narrative, we performed a two-way repeated measures analysis of

Table 1 Study 1: Manipulation Check—Targets' Self-Rated Affect in Each Affect Narrative

Self-Rating	Affect Narrative			
	Happiness	Disgust	Fear	Sadness
Happiness	7.09	2.41	2.61	1.85
Disgust	1.09	5.36	2.11	1.76
Fear	2.48	2.84	4.47	3.99
Sadness	1.18	2.64	2.69	5.88

Note. $N = 29$. Ratings were on a 1–9 scale. Congruent affective states are boldfaced.

variance (ANOVA). Targets' self-reported affect and assigned affect narrative (i.e., whether they were to narrate a sad experience, for example) were the two independent variables (with the levels of the factors being happiness, disgust, fear, and sadness for both). The test for whether targets reported experiencing each assigned affect during their narrative was the interaction of these two factors, which was highly significant, $F(9, 252) = 59.51, p < .001$. Table 1 shows that self-reported affect matched the assigned affective narrative by a large margin for all of the affects (see diagonal values compared to off-diagonal values).

Creation of Stimuli. Because of the large number of stimuli to be judged, the narrative stimuli were edited onto four different stimulus tapes. Each stimulus tape consisted of all 29 targets telling their narrative referencing the same affect, resulting in one tape for each narrated affect. Target stimuli were placed on the tapes in the same order they occurred originally so that they remained randomized, and they were separated by 15-s pauses. All stimulus tapes were low band-pass filtered using Adobe Premiere Pro CS4 so that the voices sounded somewhat muffled and the verbal content could not be deciphered. Electronic content filtering is a commonly used method of removing linguistic content from recorded spoken communication (Haskard, Williams, DiMatteo, Heritage, & Rosenthal, 2008; Rosenthal et al., 1979). This was done to prevent the judges' affect rating task from being too easy due to targets making revealing verbal statements. Thus, the judges' task was entirely about judging both visual and vocal nonverbal cues indicative of affect and personality.

Judges' Materials and Procedure. Judges made judgments of targets' Big Five traits using the Ten-Item Personality Inventory (TIPI; Gosling et al., 2003). The TIPI includes two items to judge each of the Big Five personality traits. Each item consists of two adjectives that relate to being high on the trait or two adjectives that describe being low on the trait. For example, for Neuroticism the items are (a) anxious, easily upset and (b) calm, emotionally stable. The second item is then reverse scored, and composites are created for each trait by averaging.

In order to keep the task from being too long, judges rated only eight of the 16 affects that the targets had self-rated. Judges

rated how happy, disgusted, fearful, and sad the target was, as well as how content, grossed out, anxious, and gloomy, all with a rating scale ranging from 1 (*not at all*) to 9 (*very much*).

Judges were randomly assigned to watch one of the four tapes in its entirety, rating each target's personality and affective states before moving on to the next target. Whether the judge made personality or affect ratings first was counterbalanced between subjects. There were no observed order effects, and thus order was not considered further.

Scoring of Affect Judgment Accuracy. The congruent adjective pairs rated by judges (e.g., *happy* and *content*) were highly correlated with each other and were thus averaged to create composites for happiness, disgust, fear, and sadness. For each judge, accuracy for each kind of affect was calculated as the correlation between his or her judgments of the 29 targets for that affect (e.g., disgust) and the targets' self-ratings of how much they were experiencing that affect. Additionally, a total affect judgment accuracy score was calculated for each judge by averaging his or her accuracy across the four kinds of affect. In all analyses, the Fisher- z normalizing transformation of these Pearson correlations was used for calculations (Cohen, 1988).

Scoring of Personality Judgment Accuracy. For each judge, accuracy for each Big Five personality trait was calculated as the correlation between the judge's ratings of the 29 targets on that trait (averaged across the TIPI's two items for the trait) and the targets' self-reported scores for that trait. A total personality judgment score was also calculated for each judge by averaging his or her accuracy across the Big Five traits. In all analyses, the Fisher- z normalizing transformation of these Pearson correlations was again used for calculations.

Statistical Analysis. Relations between accuracies were examined with Pearson correlations. Meta-analysis was used to understand how accuracy of judging affect was correlated with accuracy of judging personality traits. The "studies" combined for these analyses were the four different narration tapes, treated as independent replications of the affect–personality accuracy effects (independent because a different group of judges rated each narration tape). Thus, there were 20 such meta-analyses, one for each combination of affect and trait (e.g., accuracy for fear correlated with accuracy for Neuroticism), with $k = 4$ independent groups of judges in each. In addition, a corresponding meta-analysis was conducted to test the relation between total affect judgment accuracy and total personality judgment accuracy. The Comprehensive Meta-Analysis (CMA) software was used for these analyses (Borenstein, Hedges, Higgins, & Rothstein, 2005). A fixed-effects mean correlation and a combined Z -value to test whether the mean correlation exceeded zero were calculated for each combination of affect accuracy correlated with personality accuracy, as well as for total affect judgment accuracy correlated with total personality judgment accuracy.

Table 2 Studies 1 and 2: Mean Accuracy for Judging Affect and Judging Personality Traits

Accuracy	Study 1			Study 2		
	<i>M</i>	<i>SD</i>	95% CI	<i>M</i>	<i>SD</i>	95% CI
Affective state						
Happiness	.03+	.19	[.00, .07]	—	—	—
Disgust	.12***	.19	[.08, .15]	—	—	—
Fear	.08***	.18	[.05, .11]	—	—	—
Sadness	.12***	.21	[.09, .17]	—	—	—
Positive affect	—	—	—	.01	.34	[-.03, .04]
Negative affect	—	—	—	.09***	.39	[.05, .14]
Personality traits						
Neuroticism	.07***	.22	[.03, .11]	.21***	.47	[.16, .26]
Extraversion	.18***	.21	[.15, .22]	.41***	.41	[.40, .49]
Agreeableness	.05*	.21	[.01, .08]	.11***	.34	[.08, .15]
Conscientiousness	.04*	.21	[.01, .08]	.28***	.37	[.25, .33]
Openness	.09***	.23	[.05, .13]	.17***	.36	[.13, .21]

Note. Accuracy scores are correlations, per judge, between the judge's affect or personality ratings of the targets he or she viewed and the targets' criterion values for affect or personality; *p*-values refer to one-sample *t*-tests against zero. *N* = 108 to 124 in Study 1 (missing values occurred when a judge's accuracy correlation could not be calculated due to no variance across targets); *N* = 330 in Study 2. — = not applicable.

+*p* < .10. **p* < .05. ****p* < .001.

Results

Preliminary Analyses. Table 2 shows average accuracy for judging the four affect narratives and for judging the five personality traits. Accuracy was significantly above chance (i.e., greater than zero) for all accuracies except for judging happiness, for which the *t*-test was *p* < .10. One can see that absolute levels of accuracy were very modest, even while being significantly above the guessing level. Because the scoring metric was the correlation scale, in theory, accuracy could range up to 1.00. There was, fortunately, a wide range of accuracy across individuals for judging both affect and personality, meaning that the accuracy scores were acceptable candidates for prediction.

Another preliminary analysis concerned the relation of personality to self-reported affective state. For judges to have success in using affective expression as a clue to targets' personality, there should be a relation between targets' traits and their self-rated affect. Therefore, we correlated the targets' Big Five traits with how they self-rated their affective experience in the narratives. As above, the self-ratings for happiness, disgust, sadness, and fear were averaged with their three synonyms to yield one self-rating score for happiness, disgust, sadness, and fear in each affect narrative. These scores were then averaged across the four narratives.

The correlation between Neuroticism and self-rated happiness was significantly negative, $r(27) = -.48, p < .01$, meaning that targets who described themselves as more neurotic reported feeling less happy. In addition, more neurotic targets reported feeling more fear, $r(27) = .37, p = .05$. In contrast, targets' self-rated Extraversion had an inconsistent relation to their self-rated affect. Over all four narratives, the association of Extraversion with happiness was weak and actually slightly negative, $r(27) = -.07$. When individual narratives were examined, the correlation was as one would predict in the happy narrative,

with Extraversion being correlated positively with self-ratings of happiness, $r(27) = .36, p < .06$. However, in the happy narrative, Extraversion was also marginally positively correlated with sadness, $r(27) = .31, p = .10$. This inconsistency for Extraversion is consistent with the meta-analysis of Riggio and Riggio (2002), which found higher Extraversion to be associated with higher emotional expressiveness in general. Thus, whereas Neuroticism had a clear-cut relation to self-rated dysphoric emotion, Extraversion had an inconsistent relation to affective experience.

Main Analyses: Correlations Between Affect Judgment Accuracy and Personality Judgment Accuracy.

Turning to the main questions of this study, we first examined how overall accuracy for judging affect was related to overall accuracy for judging personality. This was done using the meta-analytic strategy described above. First, the correlation was calculated separately for each of the four affect narratives (disgust, fear, happiness, and sadness) that had been watched by independent groups of judges, and then they were combined using meta-analysis. The average correlation between these two overall skills was small and nonsignificant, $r = .09, p = .32$. Though not literally zero, this is a very small correlation.

To test the specific hypotheses regarding Neuroticism and Extraversion, correlations were calculated between the four affect judgment accuracy scores (for disgust, fear, happiness, and sadness) and the five personality traits (for Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience), using the same meta-analytic method. Accuracy of judging fear and accuracy of judging Neuroticism were positively related, as predicted, $r = .23, p < .05$, whereas the averaged correlation of fear accuracy with the other four traits was only $r = .09$. For Extraversion, in contrast, there was no

correlation between accuracy in judging happiness and accuracy in judging Extraversion, $r = -.01$. However, the averaged correlation of happiness accuracy with the other four trait accuracies was equally small, $r = -.02$. Thus, the predicted connection for fear and Neuroticism was found, but not for the (more weakly predicted) connection between happiness and Extraversion.

Discussion

On average, people were accurate at better than chance levels for judging both state affect and trait personality, but overall accuracies in the two domains were not correlated with each other. The central and more specific hypotheses concerned particular traits and particular affective states. The hypothesis for a connection between accuracy in judging fear and Neuroticism was supported. This finding indicates that while overall accuracies for judging traits and states may not be much connected, there is evidence that when the states and traits being judged share common cues, as in the case of distressed (e.g., fearful) affect and Neuroticism, there is a link. However, we did not find a link between accuracy in judging positive affect, specifically happiness, and accuracy of judging Extraversion. This may be because although Extraversion is generally seen as a positive trait, there was no clear evidence in our study that more extraverted targets were happier. Study 2 examined the relationship between state and trait accuracy using more general measures of positive and distressed affect rather than measuring accuracy for detecting specific affective states.

STUDY 2

Method

The complete methodology for creating the data set used in Study 2 is reported in Carney, Colvin, and Hall (2007).

Targets. Thirty (50% female) Northeastern University undergraduate students served as targets, who participated for partial fulfillment of their course requirement in introductory psychology. These 30 targets were selected from a larger group of 102 students who had been videotaped in the same task. Criteria for selecting the 30 targets were that only one individual from a given dyad was chosen, equal numbers of males and females were chosen, and the videotape contained clear audio and video.

Judges. Three hundred thirty (60% female) Northeastern undergraduate students participated as judges of the targets' behavior, and they were recruited from introductory psychology classes for partial fulfillment of their course requirement.

Target Affect Measurement. Targets' affect was measured with the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988), which consists of 20 self-rated affect adjectives.

Target Personality Measurement. Targets' personalities were measured using an average of self-report, peer report, and parent reports on the Revised NEO Personality Inventory (NEO PI-R; Costa & McCrae, 1992), a standard measure of the Big Five personality traits.

Target Procedure. While being videotaped, targets engaged in a 5-min dyadic interaction with an opposite sex partner with whom they were previously unacquainted. They were instructed to "talk about whatever you like." Target personality was measured as a composite of self-, friend, and parent ratings on the NEO PI-R (Costa & McCrae, 1992).

Creation of Stimuli. The 5-min interactions for the 30 targets were assembled onto three stimulus videotapes containing 10 targets each so that the total judging time per judge would not exceed one hour. Then, because the focus of Carney et al. (2007) was to examine judgment accuracy for video excerpts (with non-content-masked audio) of different durations and different locations within the video, multiple videotapes were created to show video clips of different durations (up to 5 min) and locations, each to be judged by a different group of judges.

Decoding Materials and Procedure. Judges were randomly assigned to one of the duration/location video conditions described above. After viewing each target's clip, they spent approximately 3 min rating characteristics of the target on a 1 (*not at all like the target person*) to 5 (*exactly like the target person*) scale. They rated targets' positive and negative affect using items from the PANAS. The PANAS items included adjectives like *enthusiastic* and *interested* for positive affect, and *scared* and *nervous* for negative affect. For the Big Five personality traits, they rated four adjectives for each of the five factors. For example, *moody*, *talkative*, *curious*, *sympathetic*, and *dependable* served, respectively, as markers for Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. This set of 20 adjectives was taken from John (1989).

Scoring of Judgment Accuracy. Accuracy was calculated for each judge, for each judged construct, across the 10 targets the judge watched using the same correlational methodology as in Study 1. Separate accuracy scores were calculated for positive and negative affect, as well as for each of the Big Five traits. Total affect and personality judgment accuracy scores were calculated by averaging accuracy scores across positive and negative affect and across the Big Five traits. In all analyses, the Fisher- z normalizing transformation of these Pearson correlations was used for calculations (and transformed back into r for presentation). All judges were considered together, regardless of the video condition to which they had been assigned.

Results

Preliminary Analyses. As in Study 1, accuracy levels were examined first. Table 2 shows average accuracy for judging

positive and negative affect and for judging the five personality traits. Accuracy was significantly above chance (i.e., greater than zero) for all accuracies except for judging positive affect, similar to the situation in Study 1 where accuracy for judging happiness was not significant. The correlation across the accuracy scores for the Big Five traits between the two studies was $r(3) = .75$, indicating good correspondence between the two studies in the relative accuracy of judging the five traits.

Also, as in Study 1, we correlated the targets' criterion scores on the Big Five traits with their self-reported affect scores in order to confirm the association between their personality and their affective experience. The correlation between their PANAS negative affect and Neuroticism was $r(28) = .60$, $p < .001$, meaning that targets who were more neurotic reported experiencing more negative affect. In contrast, targets' Extraversion was not significantly related to their positive affect, although the direction was positive as predicted, $r(28) = .20$, $p < .31$. Thus, for both studies, the link between affect and personality was much stronger for Neuroticism than for Extraversion.

Main Analyses: Correlations Between Affect Judgment Accuracy and Personality Judgment Accuracy. We first examined how overall accuracy for judging affect was related to overall accuracy for judging personality traits. This correlation was $r(328) = .30$, $p < .001$. Thus, the two overall skills had a positive relation to each other.

To test the specific hypotheses of the study, correlations were calculated between each of the two affect accuracy scores (positive and negative) and the five personality accuracy scores. Accuracy of judging Neuroticism was strongly positively correlated with accuracy of judging negative affect, $r(328) = .67$, $p < .001$, whereas the average correlation between accuracy of judging the other four personality traits and accuracy of judging negative affect was $-.01$. The correlation between accuracy of judging Extraversion and accuracy of judging positive affect was $r(328) = .29$, $p < .001$, whereas the average correlation between accuracy of judging the other four personality traits and accuracy of judging positive affect was $-.06$. Thus, the hypotheses were confirmed, showing convergent validity where predicted for Neuroticism and Extraversion, and discriminant validity for other trait–affect combinations.

Discussion

Study 2 replicated the positive relation between accuracy of judging distressed affect and accuracy of judging Neuroticism found in Study 1. In addition, Study 2 found a positive relation between accuracy of judging positive affect and accuracy of judging Extraversion, unlike Study 1. However, like Study 1, the correlation for positive affect and Extraversion was much weaker than that for negative affect and Neuroticism, consistent with the argument made above that positive affect is not a defining feature of Extraversion, though it is related to it.

META-ANALYTIC COMBINATION OF STUDIES 1 AND 2

To summarize both studies, we performed a simple meta-analytic combination, specifically the combined p -value (Stouffer method; Rosenthal, 1991). For overall affect accuracy correlated with overall personality accuracy across the two studies, the combined Z was 4.57, $p < .001$, with an unweighted average correlation between the two studies of $r = .20$.

The analogous summary for the relation of accuracy of judging distressed/negative affect with accuracy of judging Neuroticism was combined $Z = 12.28$, $p < .001$, unweighted mean $r = .48$. The corresponding summary for happiness/positive affect and Extraversion was combined $Z = 6.89$, $p < .001$, unweighted mean $r = .14$.

GENERAL DISCUSSION

In examining how accuracy of judging affect predicts accuracy of judging personality, the present research opens up novel areas for theory and research. The traditions of studying accurate affect judgment and accurate personality judgment have existed in parallel for decades, with virtually no cross-fertilization, extremely few (and only incidental) efforts to measure both kinds of accuracy, and no efforts at theoretical integration. Authors in one of these traditions rarely cite research from the other tradition, even when it is pertinent to their own concerns. However, these skills need to be studied in an integrated way because, in daily life, accurate perception of other people is an integrated process in which the perceiver is processing and judging (rightly or wrongly) a multiplicity of cues and reaching a multiplicity of judgments—if not simultaneously, then in very short order. For these two fields to make meaningful contact is a significant unfilled need. Within the personality field, there is, of course, ample recognition of the connections between emotions and traits, and we built upon this awareness in the research. Yet knowing that emotions and traits are connected within a person, and even that perceptions of emotions and traits may be connected by a perceiver, is not the same as knowing whether perceivers make *accurate* use of affect inferences when reaching *accurate* conclusions about someone's personality. Thus, the emphasis in the present research on how accuracy of affect judgment is related to accuracy of personality judgment is novel.

More specifically, the present research is theoretically grounded in two existing literatures, one on connections between emotions and traits within persons, and the other demonstrating that people do use emotional expression information when judging personality. Guided by hypothesizing about how accurate affect judgment may be related to accurate personality judgment, we reported two studies that examined accuracy of judging different dimensions of positive and negative affect, along with accuracy of judging the Big Five personality traits. An important feature of both studies is that within each study, judges made affect and personality judgments based on the same target stimuli, permitting inferences about how judges

processed those specific targets' cues. We investigated how these skills were related, both overall and for specific combinations of affect judgment accuracy and personality judgment accuracy—namely, the combinations of distressed affect and Neuroticism, and positive affect and Extraversion.

Accuracy levels for judging personality in the current studies were significantly above chance, mostly very much so, meaning that on average the judges were able to judge the traits in question, although with accuracy levels that were not high in an absolute sense. It is important to note, however, that when accuracy is scored using the present method, levels are rarely much higher than this. Connelly and Ones (2010), in their meta-analysis of over 30 studies of accuracy in judging personality at zero acquaintance that used the same correlational scoring method as we did, found overall accuracy levels in the same general range as ours and none higher than .22 (for Extraversion, which was also judged most accurately in the present studies). The correlations between Connelly and Ones's (2010) obtained accuracies and ours, across the Big Five traits, were $r(3) = .86, p = .06$, and $r(3) = .90, p < .05$, for Studies 1 and 2, respectively. Thus, the general level and the pattern of accuracy we obtained for personality accuracy was comparable to that found in the literature.

Accuracy levels in the present studies for judging affect were also low in an absolute sense, and significantly above chance only for negative affect. There is much less published data to which these figures can be compared because researchers on accuracy for judging affective states or discrete emotions rarely use correlational scoring metrics, and typically accuracy is not based on spontaneously generated expressions (as opposed to posed, prototypical expressions for which extremely high levels of judgment accuracy are the norm). Research using spontaneous expressions of emotion suggests that accuracy can be quite low although still above guessing (Hall & Schmid Mast, 2007; North, Todorov, & Osherson, 2012).

All of the results in Study 2 were stronger in terms of the magnitudes of the correlations between accuracies than in Study 1. Reasons for this can only be speculated upon, but a likely reason may be that the targets' criterion personality data in Study 2 had less random error, being a composite across several respondents rather than just the target's self-ratings. It is also possible that the trait affect measure used in Study 2 as the criterion for assessing affect judgment accuracy was a more stable and valid measure, and also less prone to random error, than the state affect measure used in Study 1. In the particular case of positive affect, a possible reason for the stronger effect in Study 2 may be that the positive items of the PANAS (used in Study 2) are more conceptually related to Extraversion (e.g., *excited, enthusiastic, proud, determined, active*) than is happiness by itself (measured in Study 1). Although extraverts do report positive affect (Augustine & Larsen, 2015; Costa & McCrae, 1980) and display positive expressions (e.g., Borkenau & Liebler, 1995), happiness is not a primary defining feature of the Extraversion construct (Back et al., 2009).

Our main finding that accuracy for judging both Neuroticism and Extraversion was related to accuracy for judging negative

and positive affect, respectively, supports a theoretical model that we call the State and Trait Accuracy Model (STAM). This model goes beyond the correlational evidence in the present studies to explicitly propose a temporal and causal ordering between these two kinds of accuracy. The STAM is based on the earlier models of Gilbert (1998) and Trope (1986, 1989; Trope & Liberman, 1993) that describe the process of person perception more broadly. In the first stage of those models, judges identify behaviors (which could include displays of affect) in terms of what has occurred or the actor's intentions, and in later stages, judges determine why the behavior was performed and make trait inferences. Rather than representing different processes, in the STAM the stages represent processing of different domains, and it is proposed that state judgments precede and inform trait judgments when no prior information is available. The first stage in STAM is similar to the models of Trope and Gilbert, in that the focus is on a state-level judgment, although the models differ in whether it is the intent of behavior in general (Gilbert's model), identification of the behavioral category (Trope's model), or judgments of affective states (STAM). The second stage of the earlier process models is attribution or dispositional inference, in which the reason for the identified behavior is determined. This stage is consistent with STAM in that information from Stage 1 is used to inform judgments made at Stage 2. Judges would make adjustments for contextual information within this model during the utilization stage of trait judgment, similar to Stage 3 of Gilbert's (1998) model.

A distinguishing feature of STAM concerns the *accuracy* of judgments. The model proposes that accurate affect judgments should facilitate accurate trait judgments because one aspect of personality is one's affect (Funder, 2013), and this order of judgments implies that affect judgments are made prior to trait judgments in a zero-acquaintance situation. This does not mean that affect is the only relevant cue of personality, and there are certainly situations in which other aspects of behavior might be better cues than affect, but STAM does contend that affect is a highly relevant cue for judgments of some personality traits, and therefore accurate judgments of affect are likely to facilitate accurate judgments of personality. This also does not mean that trait judgments are not also made very quickly, and at times quite automatically (Bar, Neta, & Linz, 2006; Todorov, Pakrashi, & Oosterhof, 2009; Uleman, Saribay, & Gonzalez, 2008), but instead predicts that trait judgments will be more accurate when they are based, at least in part, on accurate affect judgments. Because of the focus on accuracy, STAM is consistent with and builds on the Realistic Accuracy Model (RAM; Funder, 1995). The stages of RAM—relevance, availability, detection, and utilization—are incorporated within each stage of STAM (see Figure 1). In other words, STAM proposes that people first judge affective states by following the stages of RAM, and then use this information as cues to make judgments about traits, again following the stages of RAM. We propose that this process will lead to accurate trait judgments in situations of weak to moderate strength in which a person's affective reactions are guided, at least in part, by personality. This idea is

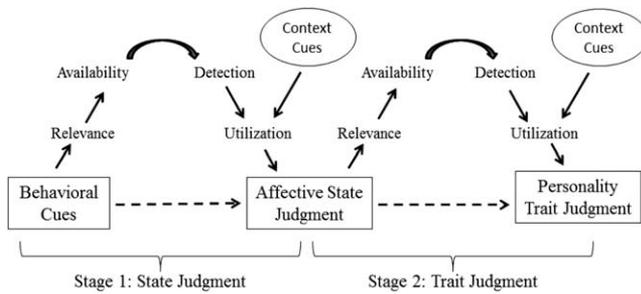


Figure 1 The State and Trait Accuracy Model.

consistent with theorizing about situational strength (Snyder & Ickes, 1985) and with Marshall and Brown's (2006) Traits as Situational Sensitivities Model.

There are a few reasons for the prediction that the causal direction is from affect judgments to trait judgments, when no prior information about the target of judgment is available. First, people use easily assessable information, including current affective states, to make judgments that generalize to the trait level, in addition to using information about categories to which people belong (which often results in stereotypical judgments). Second, affect judgments occur in part because judging affect is a lower-level process and people's implicit personality theories include affect as a component of personality. In other words, people first determine what others are like in the moment and then use that information to make judgments about more enduring aspects of personality, such as traits. Therefore, accurate judgments about current affective states are likely to lead to accurate personality judgments when affect and personality are related, and therefore affect cues are relevant to personality.

Although the present correlational studies cannot answer questions about the causal processes involved, according to the STAM, judgments about affect are first-line, lower-level judgments, which, if accurate, can inform the process of making the more complex judgment of how much a person possesses a given personality trait. Evidence supporting this also emerged in the present studies, in that targets scoring higher on Neuroticism and Extraversion tended to have concordant affective experience. An important next test of STAM will be to assess the causal link between accuracy of judging affect and accuracy of judging traits. This could be done by giving judges valid versus invalid information about affect, or valid versus invalid affective cues, and then examining how this influences the accuracy of trait judgments. Learning more about the process of accurate state and trait judgments could lead to vital breakthroughs in constructing training programs to increase the accuracy of trait judgments, which are often used to make important decisions about others, such as whom to marry or whom to hire for a job.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Note

1. For simplicity, we employ the terms *affect* and *affective states* throughout this article. In both studies, self- and other ratings of targets were made on various affective/emotional dimensions.

References

- 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.
- Augustine, A. A., & Larsen, R. J. (2015). Personality, affect, and affect regulation. In M. Mikulincer, P. R. Shaver, M. Cooper, & R. J. Larsen (Eds.), *APA handbook of personality and social psychology, Vol. 4: Personality processes and individual differences* (pp. 147–165). Washington, DC: American Psychological Association.
- Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology*, *97*, 533–548.
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*, 269–278.
- Boone, R. T., & Schlegel, K. (2016). Is there a general skill in judging others accurately? In J. A. Hall, M. Schmid Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 379–403). Cambridge: Cambridge University Press.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis* (2nd ed.). Englewood, NJ: Biostat.
- Borkenau, P., & Liebler, A. (1995). Observable attributes as manifestations and cues of personality and intelligence. *Journal of Personality*, *63*, 1–25.
- Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of Social Psychology* (Vol. 2, pp. 634–654). Cambridge, MA: Addison-Wesley.
- Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, *41*, 1054–1072.
- Clark, L. A., & Watson, D. (2008). Temperament: An organizing paradigm for trait psychology. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed.; pp. 265–286). New York: Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Connelly, B. S., & Ones, D. S. (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*, 1092–1122.
- Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*, 668–678.
- 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.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, **102**, 652–670.
- Funder, D. C. (2013). *The personality puzzle* (6th ed.). New York: W. W. Norton.
- 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.
- Gawda, B. (2007). Neuroticism, extraversion, and paralinguistic expression. *Psychological Reports*, **100**(3, Pt. 1), 721–726.
- Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vols. **1** and **2**; 4th ed.; pp. 89–150). New York: McGraw-Hill.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, **37**, 504–528.
- Hall, J. A., & Friedman, G. B. (1999). Status, gender, and nonverbal behavior: A study of structured interactions between employees of a company. *Personality and Social Psychology Bulletin*, **25**, 1082–1091.
- Hall, J. A., Gunnery, S. D., & Andrzejewski, S. A. (2011). Nonverbal emotion displays, communication modality, and the judgment of personality. *Journal of Research in Personality*, **45**, 77–83.
- Hall, J. A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion*, **7**, 438–446.
- Hareli, S., & Hess, U. (2010). What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition and Emotion*, **24**, 128–140.
- Hareli, S., Shomrat, N., & Hess, U. (2009). Emotional versus neutral expressions and perceptions of social dominance and submissiveness. *Emotion*, **9**, 378–384.
- Haskard, K. B., Williams, S. L., DiMatteo, M. R., Heritage, J., & Rosenthal, R. (2008). The provider's voice: Patient satisfaction and the content-filtered speech of nurses and physicians in primary medical care. *Journal of Nonverbal Behavior*, **32**, 1–20.
- Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2015). Accurate judgments of neuroticism at zero acquaintance: A question of relevance. *Journal of Personality*, **83**, 221–228.
- John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. M. Buss & N. Cantor (Eds.), *Personality Psychology: Recent Trends and Emerging Directions* (pp. 261–271). New York: Springer.
- Karsten, J., Penninx, B. W. J. H., Riese, H., Ormel, J., Nolen, W. A., & Hartman, C. A. (2012). The state effect of depressive and anxiety disorders on Big Five personality traits. *Journal of Psychiatric Research*, **46**, 644–650.
- Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, **20**, 165–182.
- Lucas, R. E., Le, K., & Dyrenforth, P. S. (2008). Explaining the extraversion/positive affect relation: Sociability cannot account for extraverts' greater happiness. *Journal of Personality*, **76**, 385–414.
- Marshall, M. A., & Brown, J. D. (2006). Trait aggressiveness and situational provocation: A test of the traits as social sensitivities (TASS) model. *Personality and Social Psychology Bulletin*, **32**, 1100–1113.
- McAleer, P., Todorov, A., & Belin, P. (2014). How do you say "hello"? Personality impressions from brief novel voices. *PLoS ONE*, **9**, e90779.
- Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, **27**, 237–254.
- North, M. S., Todorov, A., & Osherson, D. N. (2012). Accuracy of inferring self- and other-preferences from spontaneous facial expressions. *Journal of Nonverbal Behavior*, **36**, 227–233.
- Petrican, R., Todorov, A., & Grady, C. (2014). Personality at face value: Facial appearance predicts self and other personality judgments among strangers and spouses. *Journal of Nonverbal Behavior*, **38**, 259–277.
- Realo, A., Allik, J., Nõlvak, A., Valk, R., Ruus, T., Schmidt, M., et al. (2003). Mind-reading ability: Beliefs and performance. *Journal of Research in Personality*, **37**, 420–445.
- Riggio, H. R., & Riggio, R. E. (2002). Emotional expressiveness, extraversion, and neuroticism: A meta-analysis. *Journal of Nonverbal Behavior*, **26**, 195–218.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (revised ed.). Thousand Oaks, CA: Sage.
- Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS test*. Baltimore: Johns Hopkins University Press.
- Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or a set of modality- and emotion-specific skills? *Personality and Individual Differences*, **53**, 16–21.
- Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 883–947). Hillsdale, NJ: Erlbaum.
- Spinhoven, P., Penelo, E., de Rooij, M., Penninx, B. W., & Ormel, J. (2014). Reciprocal effects of stable and temporary components of neuroticism and affective disorders: Results of a longitudinal cohort study. *Psychological Medicine*, **44**, 337–348.
- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, **27**, 813–833.
- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, **93**, 239–257.
- Trope, Y. (1989). Levels of inference in dispositional judgment. *Social Cognition*, **7**, 296–314.
- Trope, Y., & Liberman, A. (1993). The use of trait conceptions to identify other people's behavior and to draw inferences about their personalities. *Personality and Social Psychology Bulletin*, **19**, 553–562.
- Uleman, J. S., Saribay, S. A., & Gonzalez, C. M. (2008). Spontaneous inferences, implicit impressions, and implicit theories. *Annual Review of Psychology*, **59**, 329–360.
- Vernon, P. E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology*, **4**, 42–58.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, **54**, 1063–1070.