

Impressions Based on a Portrait Predict, 1-Month Later, Impressions Following a Live Interaction

Social Psychological and
Personality Science
2017, Vol. 8(1) 36-44
© The Author(s) 2016
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1948550616662123
journals.sagepub.com/home/spp



Gul Gunaydin¹, Emre Selcuk², and Vivian Zayas³

Abstract

When it comes to person perception, does one “judge a book by its cover?” Perceivers made judgments of liking, and of personality, based on a photograph of an unknown other, and at least 1 month later, made judgments following a face-to-face interaction with the same person. Photograph-based liking judgments predicted interaction-based liking judgments, and, to a lesser extent, photograph-based personality judgments predicted interaction-based personality judgments (except for extraversion). Consistency in liking judgments (1) partly reflected behavioral confirmation (i.e., perceivers with favorable photograph-based judgments behaved more warmly toward the target during the live interaction, which elicited greater target warmth); (2) explained, at least in part, consistency in personality judgments (reflecting a halo effect); and (3) remained robust even after controlling for perceiver effects, target effects, and perceived attractiveness. These findings support the view that even after having “read a book,” one still, to some extent, judges it by its “cover.”

Keywords

impression formation, person perception, photographs, live interactions, interpersonal relationships, behavioral confirmation, halo effect, thin slices of behavior

A well-known adage urges individuals to “not judge a book by its cover.” The underlying lesson is that initial (and superficial) impressions based on an object’s outward appearance may not reflect its true worth and perceivers would revise their initial impression if afforded the opportunity to learn about the true nature of the object. In contrast, an idea dating back to the Ancient Greeks is that the face reveals aspects of personality. This view suggests that outward appearance provides reliable information about inner character and thus perceivers’ initial impressions would remain relatively unchanged even after having the opportunity to learn additional information about the object. When it comes to person perception, which view is supported?

Lay intuitions reflect the “do not judge a book by its cover” idiom: People overwhelmingly hold the belief that they are revisionists. In an independent study ($N = 114$; Study S1), people believed that their initial impressions of another person based on a portrait photograph would be revised if they had the opportunity to actually interact with the person. Notwithstanding these intuitions, to date, there has been no direct empirical investigation of this age-old issue. The lack of attention is surprising; there has been considerable research on both impressions formed based on viewing static photographs (Gunaydin, Zayas, Selcuk, & Hazan, 2012; Shoda & McConnell, 2013; Tabak & Zayas, 2012; Willis & Todorov, 2006; Wilson & Rule, 2015) and those formed following real-life interactions (Kraus & Keltner, 2009; Mendes & Koslov, 2013; Reis et al.,

2010). But, by in large, there has been little cross-fertilization between these two lines of work.

A handful of studies have investigated the extent to which judgments of attractiveness based on a target’s photograph predict judgments of attractiveness based on viewing a video clip of the same target reciting scripted text (Lander, 2008; Rubenstein, 2005) or describing an autobiographical memory (Roberts et al., 2009). However, because video clips are one sided (vs. bidirectional) and restrict exposure to interpersonal cues (Aviezer, Trope, & Todorov, 2012; Tickle-Degnen & Rosenthal, 1990), they do not offer the same phenomenological experience as a real-life, ongoing, dyadic interaction (Zayas, Shoda, & Ayduk, 2002).

Also relevant is research on behavioral confirmation (or self-fulfilling prophecy) in which perceivers’ initial judgments influence how they behave toward the target, which in turn elicits behaviors from the target that are consistent with the initial judgment (Snyder & Stukas, 1999; Snyder, Tanke, &

¹ Bilkent University, Ankara, Turkey

² Middle East Technical University, Ankara, Turkey

³ Cornell University, Ithaca, NY, USA

Corresponding Author:

Gul Gunaydin, Bilkent University, Ankara 06800, Turkey.
Email: gulgunaydin@bilkent.edu.tr

Berscheid, 1977). In seminal research by Snyder, Tanke, and Berscheid (1977), male perceivers were shown a photograph of either an attractive or an unattractive woman. Men who believed they were talking on the phone to an attractive (vs. unattractive) woman behaved more warmly toward her, and this in turn elicited warmer behaviors from her (all behaviors were judged by independent coders). Despite the relevance of this work, it did not test whether photograph-based judgments predict judgments following a face-to-face interaction. Because there was no face-to-face interaction, perceivers were blind to the target's appearance and other visual information that could serve to color in vivo impressions. Thus, the question remains: Do judgments based on a portrait photograph predict judgments following an actual in-person interaction?

Despite the lack of direct empirical investigations, a body of research on physical appearance supports a *consistency hypothesis*. Judgments based on the face occur spontaneously, effortlessly, and sometimes outside of conscious awareness (Gunaydin et al., 2012; Willis & Todorov, 2006). Once perceivers form favorable impressions of another person they tend to attribute desirable characteristics to them (Nisbett & Wilson, 1977), and in an actual interaction possibly treat them in ways that confirm initial favorable impressions (Snyder et al., 1977). Thus, to the extent that facial cues color both photograph- and interaction-based judgments, we would predict consistency between the two.

On the other hand, research also supports a *revision hypothesis*. Actual interactions are not only influenced by facial appearance but an array of other cues conveyed through the body (Aviezer et al., 2012), including nonverbal behaviors (Tickle-Degnen & Rosenthal, 1990), tone of voice (Berry, 1992), speech content (Ireland et al., 2011), and clothing (Mills & Aronson, 1965). Live interactions also afford opportunities to learn about a person and observe the person's behaviors (Letzring, Wells, & Funder, 2006).

Present Research

The goal of the present research was to investigate the extent to which a consistency (vs. a revision) hypothesis is supported. We asked perceivers to make judgments of liking, and of personality, of four unknown others (targets) based solely on their portrait photograph (consisting of the face and hair). At least 1 month later, perceivers had an extensive one-on-one interaction with one of these targets across two different contexts: a 10-min structured interaction (trivia game) that afforded limited opportunities to learn about each other and a 10-min unstructured interaction (getting-to-know interaction) in which both parties tried to get to know each other as much as they could. This design offered several advantages. First, it allowed us to examine our primary question of whether, and to what extent, photograph-based judgments predict interaction-based judgments (hereafter referred to as the *photograph-live association*). Given the only information available in both portrait photographs and live interactions is the facial appearance of the target, a statistically significant photograph-live association would provide evidence that judgments cued by facial

appearance influence judgments in a live interaction. Second, by assessing the photograph-live association during an extensive 20-min social interaction that involved a structured setting in which the exchange of personal information is limited and an unstructured setting in which individuals are encouraged to get to know each other, this design allowed us to examine the extent to which judgments might be revised with greater interpersonal knowledge (Letzring et al., 2006).

We predicted that with respect to liking judgments, which are closely linked with trustworthiness judgments, the results would favor the consistency hypothesis. Judgments of trustworthiness, partly because of their functional importance from an evolutionary standpoint (Cosmides & Tooby, 1992), are among those formed the fastest (Willis & Todorov, 2006) and are difficult to override (Zajonc, 1980). To go beyond initial impressions based on the face, perceivers need to observe behaviors diagnostic of whether or not a person can really be trusted (e.g., If I told this person a secret, would they keep it?). Based on theories of interpersonal trust, such diagnostic behaviors may not reveal themselves in an initial encounter (Simpson, 2007), in which case perceivers may rely more heavily on their initial impressions of liking. This would be reflected in a statistically significant photograph-live association in liking that would be observed in both types of interactions (trivia vs. getting-to-know; Hypothesis 1).

We also predicted consistency with regard to personality judgments, but to a lesser extent and for different reasons. Based on the literature on the halo effect by which perceivers ascribe socially desirable personality characteristics (Backstrom, Bjorklund, & Larsson, 2009) to targets that they like, we reasoned that personality judgments would be colored by liking judgments (Nisbett & Wilson, 1977) and, thus, show some consistency. However, whereas the basis of liking judgments is not clearly known, the basis for personality judgments is more easily seen (e.g., orderliness cues conscientiousness), and because there are observable behaviors, personality judgments may be more likely to be revised (Funder, 2012; Vazire, 2010). Thus, we predicted a (more modest) photograph-live association for personality judgments (Hypothesis 2).

Finally, we explored potential mechanisms for consistency between photograph- and interaction-based liking judgments. In particular, we looked for evidence that the photograph-live association for liking judgments would emerge, in part, due to behavioral confirmation (Snyder et al., 1977). That is, we examined whether perceivers' favorable photograph-based impressions would predict their *actual behaviors* (e.g., warmth) toward the target, which in turn would elicit behaviors from the target that confirm initial impressions (Darley & Fazio, 1980; Snyder & Stukas, 1999).

Method

Participants

Six hundred and thirty-six adults (417 females, mean age = 20.65 years, $SD = 2.65$ years) completed the Time 1 measures

in an online survey. We invited all individuals who met the pre-determined selection criteria (see Procedures and Measures for details) to the Time 2 lab session and aimed to run all participants who accepted the invitation within two academic semesters. We stopped data collection at the end of two consecutive semesters. Of the 348 individuals invited to the lab session, 56 accepted the invitation. One participant was excluded from the final analyses because she expressed after the lab session that she already knew the target. Therefore, the Time 2 sample consisted of 55 participants (37 females, mean age = 19.93 years, $SD = 1.94$ years). With the current sample, statistical power ($1 - \beta$) for detecting an average (for the sample) within-person standardized association of .3 between photograph- and interaction-based judgments was 91% and the power to detect a difference of .3 in photograph–live association across interaction contexts (trivia minus getting-to-know) was 78% (see Supplemental Materials for calculation of statistical power).

Procedures

All instructions, measures, and the data may be downloaded from the Open Science Framework (URL: https://osf.io/nhmtw/?view_only=9f6efafeba4b48dc9b6a73b6a3d145ee). Descriptives and reliabilities of the measures as well as further details about the procedures are reported in the Supplemental Materials.

Time 1. In an online survey, participants viewed a total of eight portrait photographs: photographs of four female targets, each displaying a smile or a neutral expression. Participants were first asked whether or not they personally knew any of the targets. For those targets they indicated *not* knowing, perceivers were asked to provide their impressions in two separate blocks, presented in counterbalanced order. In one block, targets displayed a smiling expression in the photograph and in another block they displayed a neutral expression. Within each block, the order of the target was randomly determined. Questionnaires not relevant to the present study and demographic questions separated the blocks containing photographs with different emotional displays.

After viewing each photograph, perceivers were asked to evaluate the person in the photograph on liking, attractiveness, and personality. Half of the perceivers completed the personality items first, and the other half completed the liking and attractiveness items first. We computed *photograph-based judgments* for each target by averaging across judgments of the neutral and smiling photograph. In this way, photograph-based judgments reflect more context-independent evaluations (Todorov & Porter, 2014).

Time 1 photograph-based judgments were used to select a subset of individuals to participate in the lab session (i.e., hereafter referred to as *perceivers*). Perceivers were selected, so that each of the four targets interacted with roughly equal number of perceivers who held favorable photograph-based impressions of her and who held unfavorable impressions.¹ In this

way, we aimed to equate the likability of the targets, minimizing the likelihood that associations between photograph- and interaction-based judgments would simply be driven by *target effects* (i.e., perceivers evaluating some targets more favorably than others). Moreover, perceivers who did not discriminate among different targets in their evaluations (i.e., perceivers evaluating all targets favorably or those evaluating all targets unfavorably) were not invited to the lab session. This would serve to minimize *perceiver effects* (i.e., the tendency to like or dislike others in general; see Supplemental Materials for further information about participant selection).

Time 2. At least 1 month ($M = 4.22$ months, $SD = 2.27$ months) after completing the Time 1 survey, perceivers participated in a study on “social interactions.” Perceivers were told that they would interact with another participant while being videotaped. In fact, unbeknownst to them, the other participant was one of the targets whose photographs they had evaluated at Time 1. Each perceiver–target pair completed two interactions adapted from a well-established paradigm (Letzring et al., 2006). The first interaction was a 10-min “trivia game” that consisted of 20 questions of low, moderate, and high difficulty taken from a database of questions normed for difficulty (Nelson & Narens, 1980). Each pair was asked to work on the questions until their time was up. The second interaction was a 10-min “getting-to-know interaction” in which the perceiver–target pair was instructed to get to know each other as well as possible. The dyads were asked to complete the more superficial trivia game interaction first followed by the more intimate getting-to-know interaction for two reasons: (1) Possible effects of the two interactions on each other are asymmetrical. Personal information learned through the getting-to-know interaction cannot be undone during the trivia game whereas the little information obtained through the structured trivia game is less likely to have an effect on the getting-to-know interaction. (2) Day-to-day interactions often transition gradually from superficial to intimate (Aron, Melinat, Aron, Vallone, & Bator, 1997). At the end of each 10-min interaction, the target left the room, and perceivers and targets provided impressions of each other based on the interaction using the same measures as those administered at Time 1.

Assessing the Photograph–Live Association in Liking

At both time points, perceivers were asked to indicate liking of the target on four traits used in past work tapping trustworthiness/warmth (*trustworthy, likable*) and dominance/competence (*competent, aggressive*; Willis & Todorov, 2006) on a 7-point scale (1 = *not at all* to 7 = *extremely*) and 4 items assessing willingness to approach and engage in future interactions with the person (e.g., “*This seems like the kind of person whom I would like to get to know*”; 1 = *strongly disagree* to 7 = *strongly agree*). After reverse scoring aggressiveness, these 8 items were averaged to index liking of the target based on the photograph as well as following the trivia game and the getting-to-know interaction. These indices were highly reliable.

Assessing the Photograph–Live Association in Personality

At both time points, perceivers indicated their impressions of each target’s personality using items adapted from a well-validated and widely employed instrument (Gosling, Rentfrow, & Swann, 2003) measuring five major personality dimensions—extraversion, agreeableness, neuroticism, conscientiousness, and openness to experience. The possible range of scores for the bipolar personality scales was 1 (e.g., *reserved, quiet*) to 7 (e.g., *extraverted, enthusiastic*).

Assessing Behavioral Confirmation: Coding Thin Slices of Behavior

We asked eight independent coders (six females) who were blind to the study hypotheses and perceivers’ photograph-based liking to watch (without audio) 30-s thin slices (Ambady & Rosenthal, 1993) selected from the start, middle, and end of each interaction, with the slices beginning, respectively, 30 s, 4.5 min, and 8.5 min after the dyad started conversing during a given interaction. For each slice, the coders rated the person in the video in terms of warmth (“How warm was this person being toward their interaction partner?”), “How much do you think this person enjoyed the interaction?”, and “How engaged was this person during the interaction?”; 1 = *not at all* to 7 = *very much*). Importantly, given the perceiver and the target were each captured by a different video camera, coders rated each person without seeing the behaviors of the interaction partner. For each interaction, we computed a behavioral warmth score by averaging across evaluations based on the start, middle, and end of the interaction.

Data Analytic Strategy

Because perceivers’ judgments were nested within targets, to test Hypotheses 1 and 2, we used linear mixed models (LMMs) to account for interdependency among data points when estimating the photograph–live association in liking and personality judgments. Conceptually, this means we estimated the photograph–live association in liking for each target and then averaged the photograph–live association across the four targets. Critically, because the photograph–live association in liking is estimated for each target separately and then combined, any observed consistency cannot be attributed simply to differences in target characteristics (*target effects*) in which some targets are generally more liked than others.

To compare the size of the photograph–live associations across interactions (trivia vs. getting-to-know) and type of judgment (liking vs. personality), we constructed 95% confidence intervals (CI) around the difference of the two critical associations (e.g., photograph–live association in liking minus photograph–live association in personality) using the modified asymptotic method (Zou, 2007). If the CI did not include zero, we concluded that the two associations were significantly different from each other. To investigate whether consistency in liking judgments accounts for consistency in personality judgments (or vice versa),

we conducted path analyses using LISREL v. 9.1. Finally, we investigated potential mediating role of behavioral confirmation using a joint significance test, an approach that has higher statistical power compared to traditional approaches (e.g., product of coefficients methods) while minimizing Type I error (Taylor, MacKinnon, & Tein, 2008).

Results

Does Photograph-Based Liking Predict Interaction-Based Liking?

Providing evidence of some consistency, photograph-based judgments of liking predicted 35% of the variance in interaction-based judgments of liking ($B = .489$, $SE = .063$, $p < .001$; 95% CI [0.363, 0.615]; Figure 1).² The photograph–live association for liking judgments did not differ significantly as a function of interaction type (95% CI [−0.093, 0.190]).³ Although our statistical power to detect a medium-sized difference in photograph–live association across interaction contexts was good (78%), this null finding should be interpreted with caution. But importantly, we observed a significant photograph–live association both following the trivia game and the getting-to-know interaction (Table 1).

Accounting for the Role of Target and Perceiver Effects

Could the photograph–live association in liking be due to differences between targets (*a target effect*)? Perhaps some targets were more attractive or more likable than others. Could this difference between targets have contributed, or somehow affected, the observed results? As shown in Figure 1, this was not the case. The photograph–live association was clearly observed for each of the four targets. For example, perceivers who liked Confederate 1 based on the photograph were more likely to like Confederate 1 based on the interaction, whereas those who disliked Confederate 1 based on the photograph were more likely to hold less favorable impressions following the interaction. Because perceivers were nested within confederates, the analysis holds the effect of confederate constant.

Further, we investigated whether the photograph–live association in liking varied by target by conducting a separate study ($N = 48$) using a larger number of targets (14 confederates). In a laboratory session, perceivers were first presented with a smiling photograph of the target and asked to make liking judgments. Immediately following, they played a 10-min trivia game with the same target and again made liking judgments (see Study S2 in Supplemental Materials for further details about the procedures). Consistent with the findings here, the photograph–live association was significant ($B = .745$, $SE = .159$, $p < .001$, 95% CI [0.425, 1.064], *variance explained* = 31%), suggesting that this association is not driven by targets effects.

Alternatively, could the photograph–live association be due simply to differences in perceiver characteristics (*a perceiver effect*)? Perhaps some perceivers are more likely to express favorable impressions about unknown others; could such

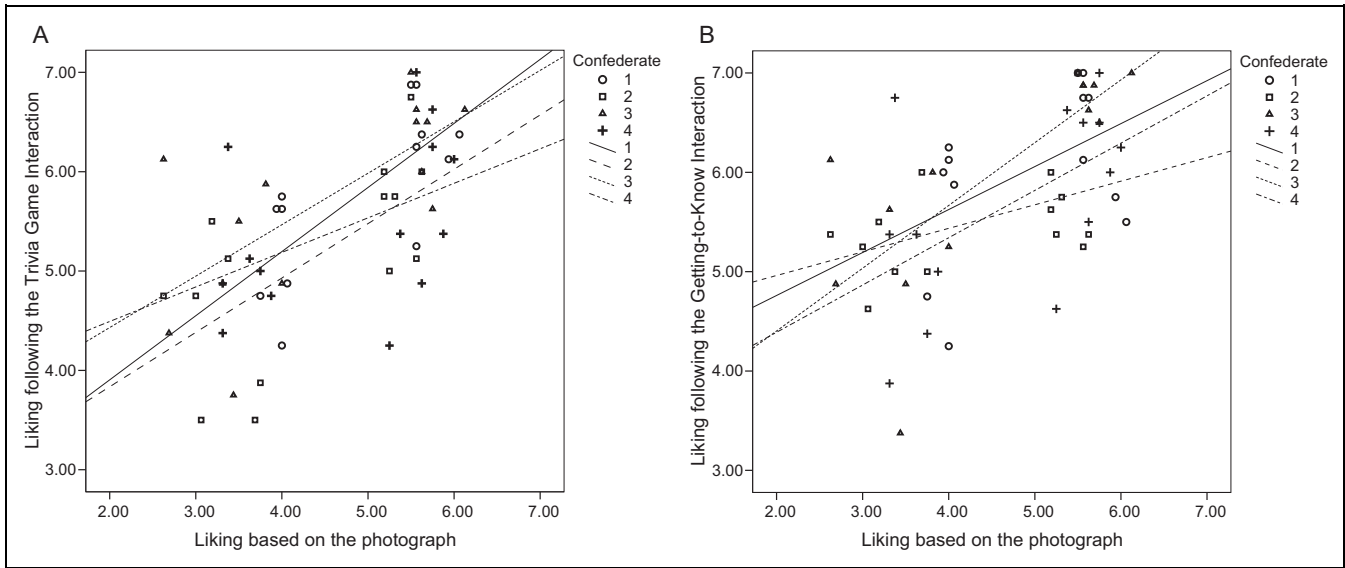


Figure 1. Scatter plots showing the association between liking judgments based on photographs (x-axis) and those following the trivia game interaction (y-axis; A) and those following the interaction aimed at getting to know the other person (y-axis; B). In each panel, each data point represents a perceiver’s judgment about the particular target with whom they interacted. The photograph–live association was computed for each of the four targets separately. The four different markers represent perceivers’ judgments of the four different targets, and the different lines show the photograph–live association for the four different targets.

Table 1. Photograph–Live Association in Liking and Personality (Agreeableness, Neuroticism, Openness, Conscientiousness, and Extraversion) Judgments as a Function of the Interaction Type (Trivia Game vs. Getting-to-know).

Judgment	Live Interaction	B	SE	p	Variance Explained	95% Confidence Intervals (CI)
Liking	Trivia game	.517	.088	<.001	38%	[0.339, 0.695]
	Getting-to-know	.462	.089	<.001	33%	[0.284, 0.640]
Agreeableness	Trivia game	.506	.161	.003	9%	[0.183, 0.830]
	Getting-to-know	.403	.126	.002	13%	[0.150, 0.656]
Neuroticism	Trivia game	.466	.156	.004	9%	[0.153, 0.779]
	Getting-to-know	.257	.149	.092	2%	[−0.043, 0.556]
Openness	Trivia game	.474	.156	.004	13%	[0.162, 0.786]
	Getting-to-know	.320	.167	.061	6%	[−0.015, 0.655]
Conscientiousness	Trivia game	.565	.126	<.001	26%	[0.313, 0.817]
	Getting-to-know	.243	.124	.056	5%	[−0.006, 0.493]
Extraversion	Trivia game	.074	.171	.665	0%	[−.269, .418]
	Getting-to-know	.273	.143	.062	0.3%	[−0.014, 0.561]

Note. Variance explained by a predictor was estimated by computing the proportional decrease in the residual variance of the model as a result of adding the predictor into the model.

differences have affected the results? To address this account, we computed a *perceiver dispositional liking* score by averaging across each perceiver’s liking of the three targets evaluated at Time 1, but with whom they did *not* interact at Time 2. In addition, to further control for possible target effects, we calculated a *target consensual likability* score for each of the four targets by averaging across photograph-based liking scores of the 602 respondents who rated all target photographs but did not participate in the lab session. Then, in our model predicting interaction-based liking, we added as covariates perceiver dispositional liking and target consensual likability scores, along with the target’s photograph-based attractiveness as rated by the perceiver. Together these three variables explained 37%

of variance in interaction-based liking (perceiver dispositional liking, $B = .530$, $SE = .137$, $p < .001$, 95% CI [0.258, 0.802]; target consensual likability, $B = 1.344$, $SE = .445$, $p = .003$, 95% CI [0.461, 2.226]; perceiver’s photograph-based attractiveness judgments, $B = .212$, $SE = .049$, $p < .001$, 95% CI [0.116, 0.309]). Critically, even after statistically controlling for these factors, the photograph–live association remained statistically significant ($B = .266$, $SE = .085$, $p = .002$, 95% CI [0.097, 0.434], *variance explained* = 8%). These results indicate that the photograph–live association is not fully explained by a general tendency for perceivers to like (vs. dislike) unknown others and that it reflects more than perceived attractiveness of the targets and provide further evidence that the

Table 2. Comparison of Photograph–Live Association in Liking Versus Personality Judgments.

Judgments	Trivia Game	Getting-to-Know
Liking vs. agreeableness	[−0.022, 0.447]	[−0.058, 0.385]
Liking vs. neuroticism	[−0.004, 0.482]	[0.087, 0.583]
Liking vs. openness	[−0.032, 0.516]	[0.031, 0.620]
Liking vs. conscientiousness	[−0.129, 0.335]	[0.047, 0.596]

Note. The size of the photograph–live associations for liking versus personality judgments was compared by constructing 95% confidence intervals around the difference of the two associations obtained from the same sample (photograph–live association in liking minus photograph–live association in personality). If the confidence interval did not include zero, the two associations were significantly different from each other. Consistency estimates were similar for liking vs. personality judgments after the trivia game interaction. However, consistency estimates were weaker for conscientiousness, neuroticism, and openness (but not agreeableness) judgments as compared with liking judgments after the getting-to-know interaction. Confidence intervals indicating significant differences are in bold font.

photograph–live association does not simply reflect a tendency for certain targets to elicit greater liking.

Do Photograph-Based Personality Judgments Predict Interaction-Based Personality Judgments?

With regard to personality judgments, for four (of five) major personality traits (agreeableness, conscientiousness, openness to experience, and neuroticism), photograph-based judgments predicted interaction-based judgments (Table 1).⁴ However, personality judgments of conscientiousness, neuroticism, and openness showed greater revision, compared to liking judgments, suggesting that perceivers more readily use individuating information to revise personality judgments than liking judgments (Table 2). Additionally, photograph-based judgments did not predict interaction-based judgments for extraversion. Given that extraversion is easily inferred from observable behaviors such as talkativeness (Vazire, 2010), perceivers might have relied less on facial appearance to judge this trait during live interactions.

Interestingly, across the four traits, consistency in personality judgments was weaker after the getting-to-know interaction as compared to the trivia interaction (Table 1). However, the decrease in consistency was statistically significant only for conscientiousness but not so for agreeableness, neuroticism, or openness (Table S1). Although our statistical power to detect a medium-sized difference in photograph–live association across interaction contexts was good (78%), these lack of statistically significant differences should be interpreted with caution.

What Accounts for Consistency in Liking Judgments? Behavioral Confirmation Dynamics

A closer look at the data indicated that the photograph–live association in liking emerged, at least in part, as a result of behavioral confirmation. We performed mediation analyses using the independent coders' assessments of perceivers' and targets' behaviors. Perceivers' photograph-based liking

predicted their behavioral warmth toward the target during the interaction ($B = .161$, $SE = .067$, $p = .017$, 95% CI [0.029, 0.293]), which in turn predicted the target's behavioral warmth toward the perceiver ($B = .375$, $SE = .047$, $p < .001$, 95% CI [0.282, 0.468]), and completing the cycle, target's behavioral warmth predicted the perceiver's interaction-based liking ($B = .380$, $SE = .119$, $p = .002$, 95% CI [0.144, 0.617]). This three-path indirect association accounted for 5% of the total photograph–live association in liking, reflecting a nontrivial, small-to-moderate indirect association (see Supplemental Materials for further details of the mediational analyses).

What Accounts for Consistency in Personality Judgments? A Halo Effect Mechanism

Research suggests that people often use liking judgments to inform personality judgments, especially in situations with minimal information. We therefore reasoned that consistency in personality judgments may be due to consistency in liking judgments. We tested this possibility using path analyses for the four personality traits (agreeableness, neuroticism, conscientiousness, and openness) that showed statistically significant photograph–live associations (Model 1 in Figure S2), and compared it to an alternative account that personality judgments explain the consistency in liking judgments (Model 2 in Figure S2). For all four personality traits, Model 1 provided a better fit than Model 2 (see Supplemental Materials for statistics). Moreover, after statistically controlling for interaction-based liking, the photograph–live association for personality judgments became weaker, indicating that interaction-based liking judgments at least partially mediated consistency in personality judgments.

Ruling Out Awareness as an Alternative Interpretation

Could the photograph–live association be driven by perceivers' motivation to maintain consistency between their photograph- and interaction-based judgments? This possibility is unlikely: The two judgments were separated by at least 1 month (range = 35–212 days) and the majority (89%) of perceivers were unaware that they were interacting with someone whose photograph they had seen earlier. Four perceivers reported during debriefing having seen the photograph of the target in the online survey and two reported that the target looked familiar. When we repeated the analyses excluding these perceivers, the photograph–live association in liking ($B = .498$, $SE = .067$, $p < .001$, 95% CI [0.365, 0.632], *variance explained* = 36%) and four personality judgments (agreeableness, neuroticism, conscientiousness, and openness; all $t_s > 2.775$, all $p_s < .007$, *variance explained* ranging from 5% to 15%) remained statistically significant.

Discussion

Lay intuitions strongly adhere to the idiom to not judge a book by its cover (Study S1). Yet, our findings provide compelling evidence that perceivers continue to be influenced by outward appearances even after interacting with a target face-to-face.

Specifically, judgments based only on a portrait photograph predicted, at least 1 month later, judgments based on an actual in-person interaction. We observed consistency for both liking judgments and to a lesser extent, personality judgments, and even when there were ample opportunities to get to know the person.

Evidence of Consistency and Revision

Even though judgments were made across *different interaction contexts* and *at least a month apart*, judgments based on the portrait photograph predicted approximately 35% of the variance in judgments following the live 20-min interaction. This finding suggests both robust consistency, but also revision. With regard to consistency, as a point of comparison, test–retest correlations for photograph-based attractiveness judgments across a *1 week* time period range between .72 and .74 (Honekopp, 2006). In the current research, the approximate photograph–live correlation is .59, suggesting appreciable consistency. Nonetheless, a considerable amount of variance remains unexplained, suggesting that initial judgments are not perfect predictors of judgments following an actual interaction.

Not All Judgments Are Alike

We observed consistency in judgments of four of the five personality traits. Interestingly, personality judgments showed evidence of greater revision than liking judgments, consistent with diverse literatures (Funder, 2012; Simpson, 2007). Liking in particular is closely linked with trustworthiness, which from an evolutionary perspective is theorized to serve a functional role in survival and reproduction (Cosmides & Tooby, 1992). Indeed, it is among the judgments made most quickly about another (Willis & Todorov, 2006). Because actual behavioral cues about trustworthiness are not easily observed (Simpson, 2007), perceivers may rely on initial impressions formed spontaneously from facial appearance. The functional role of personality judgments on the other hand is less clear. Moreover, such judgments are more readily informed by actual behaviors. In our work, judgments of extraversion showed considerable revision following an actual interaction, which may be because behaviors that cue the trait, such as talkativeness, are easily observed (Vazire, 2010).

The Source of Consistency

Why do perceivers in our present study show such consistency? With regard to liking judgments, our findings indicate that behavioral confirmation partially accounted for the photograph–live association. Perceivers with favorable photograph-based impressions of the target also acted warmly toward her during the interaction. Perceivers' warmth in turn elicited greater warmth from the target, which in turn led to more favorable interaction-based impressions. Given facial appearance is processed spontaneously, effortlessly, and sometimes outside of conscious awareness (Gunaydin et al., 2012; Willis &

Todorov, 2006), it is possible that at the very first moments of the interaction perceivers formed impressions based on the target's facial appearance, which in turn led them to behave in ways that confirm initial impressions. Critically, unlike previous demonstrations of behavioral confirmation (Snyder et al., 1977), the current work shows evidence of the powerful effect of preexisting judgments on subsequent interactions when the initial judgments are self-generated (vs. experimentally manipulated), idiosyncratic (vs. consensually agreed upon), more subtle (vs. obvious), and not simply reflecting a strong and pervasive stereotype that “beautiful is good.”

With regard to consistency in personality judgments, our results indicate the role of a halo effect. Given the primacy of liking judgments in person perception, initial favorable opinions cued by facial appearance colored more objective judgments of personality. Thus, the consistency in liking judgments gave rise, at least in part, to consistency in personality judgments.

Future Directions

Under what circumstances, and for whom, might we observe greater revision of impressions based on facial appearance? Past research suggests that perceivers might override initial impressions to a greater extent if sufficiently motivated to do so (e.g., because they expect to work with the target later) or if simply instructed to try to form accurate impressions (Neuberg & Fiske, 1987). The ability and willingness to revise may depend on individual differences, especially gender which has been found to predict accuracy in interpersonal judgments (Chan, Rogers, Parisotto, & Biesanz, 2011; Hall, 1978; Vogt & Colvin, 2003). Greater revision might also occur if perceivers discover highly negative information about a target (e.g., that the person is a child molester; Cone & Ferguson, 2015). Still, given that in everyday social interactions individuals are motivated to paint a positive picture of themselves, as well as to selectively look for positive information in others, such negative information is unlikely to be encountered when we meet a new person—for example, when introduced to a new friend or a potential business partner, or interview for a job.

Summary

Even when perceivers try to get to know unknown others as much as possible—as is common in many interview settings—they do not completely revise their initial impressions and rely on facial appearance to inform interpersonal judgments. In other words, even after having read a book, we still, to some extent, judge it by its cover.

Declaration of Conflicting Interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded in part by Cornell's Institute of Social Sciences.

Supplemental Material

The online data supplements are available at <http://journals.sagepub.com/doi/suppl/10.1177/1948550616662123>.

Notes

1. Liking status based on the photograph judgment (liking vs. disliking the target) did not moderate the photograph–live association in liking ($B = -.625$, $SE = .463$, $p = .180$) or in personality ($B = .047$, $SE = .202$, $p = .818$ for extraversion; $B = -.174$, $SE = .256$, $p = .498$ for agreeableness; $B = -.356$, $SE = .204$, $p = .084$ for conscientiousness; $B = -.358$, $SE = .241$, $p = .140$ for neuroticism; $B = -.304$, $SE = .235$, $p = .200$ for openness).
2. The photograph–live association in liking was evident for both genders. However, the association was stronger for females (see Supplemental Materials).
3. Another way of testing whether the photograph–live association depends on the interaction type is to enter the main effect of interaction type (trivia vs. getting-to-know) and its two-way interaction with the photograph-based judgment into the linear mixed models predicting interaction-based judgment. Analyses based on this approach revealed that interaction type did not moderate the photograph–live association for any of the judgments, all $t_s < 1.819$, all $p_s > .071$.
4. Although females exhibited a stronger photograph–live association in judgments of neuroticism, gender did not significantly moderate the photograph–live association in judgments of agreeableness, openness, and conscientiousness (see Supplemental Materials).

References

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*, 431–441. doi:10.1037/0022-3514.64.3.431
- Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, *23*, 363–377. doi:10.1177/0146167297234003
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, *338*, 1225–1229. doi:10.1126/science.1224313
- Backstrom, M., Bjorklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, *43*, 335–344. doi:10.1016/j.jrp.2008.12.013
- Berry, D. S. (1992). Vocal types and stereotypes: Joint effects of vocal attractiveness and vocal maturity on person perception. *Journal of Nonverbal Behavior*, *16*, 41–54. doi:10.1007/bf00986878
- Chan, M., Rogers, K. H., Parisotto, K. L., & Biesanz, J. C. (2011). Forming first impressions: The role of gender and normative accuracy in personality perception. *Journal of Research in Personality*, *45*, 117–120. doi:10.1016/j.jrp.2010.11.001
- Cone, J., & Ferguson, M. J. (2015). He did what? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology*, *108*, 37–57. doi:10.1037/pspa0000014
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York, NY: Oxford University Press.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist*, *35*, 867–881. doi:10.1037/0003-066X.35.10.867
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*, 177–182. doi:10.1177/0963721412445309
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*, 504–528. doi:10.1016/s0092-6566(03)00046-1
- Gunaydin, G., Zayas, V., Selcuk, E., & Hazan, C. (2012). I like you but I don't know why: Objective facial resemblance to significant others influences snap judgments. *Journal of Experimental Social Psychology*, *48*, 350–353. doi:10.1016/j.jesp.2011.06.001
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, *85*, 845–857. doi:10.1037/0033-2909.85.4.845
- Honekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology-Human Perception and Performance*, *32*, 199–209. doi:10.1037/0096-1523.32.2.199
- Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, *22*, 39–44. doi:10.1177/0956797610392928
- Kraus, M. W., & Keltner, D. (2009). Signs of socioeconomic status: A thin-slicing approach. *Psychological Science*, *20*, 99–106. doi:10.1111/j.1467-9280.2008.02251.x
- Lander, K. (2008). Relating visual and vocal attractiveness for moving and static faces. *Animal Behaviour*, *75*, 817–822. doi:10.1016/j.anbehav.2007.07.001
- Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *91*, 111–123. doi:10.1037/0022-3514.91.1.111
- Mendes, W. B., & Koslov, K. (2013). Brittle smiles: Positive biases toward stigmatized and outgroup targets. *Journal of Experimental Psychology: General*, *142*, 923–933. doi:10.1037/a0029663
- Mills, J., & Aronson, E. (1965). Opinion change as a function of the communicator's attractiveness and desire to influence. *Journal of Personality and Social Psychology*, *1*, 173–177. doi:10.1037/h0021646
- Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowing ratings. *Journal of Verbal Learning and Verbal Behavior*, *19*, 338–368. doi:10.1016/s0022-5371(80)90266-2

- Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53*, 431–444. doi:10.1037/0022-3514.53.3.431
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*, 250–256. doi:10.1037/0022-3514.35.4.250
- Reis, H. T., Smith, S. M., Carmichael, C. L., Caprariello, P. A., Tsai, F.-F., Rodrigues, A., & Maniaci, M. R. (2010). Are you happy for me? How sharing positive events with others provides personal and interpersonal benefits. *Journal of Personality and Social Psychology*, *99*, 311–329. doi:10.1037/a0018344
- Roberts, S. C., Saxton, T. K., Murray, A. K., Burriss, R. P., Rowland, H. M., & Little, A. C. (2009). Static and dynamic facial images cue similar attractiveness judgements. *Ethology*, *115*, 588–595. doi:10.1111/j.1439-0310.2009.01640.x
- Rubenstein, A. J. (2005). Variation in perceived attractiveness—Differences between dynamic and static faces. *Psychological Science*, *16*, 759–762. doi:10.1111/j.1467-9280.2005.01610.x
- Shoda, T. M., & McConnell, A. R. (2013). Interpersonal sensitivity and self-knowledge: Those chronic for trustworthiness are more accurate at detecting it in others. *Journal of Experimental Social Psychology*, *49*, 440–443. doi:10.1016/j.jesp.2012.12.011
- Simpson, J. A. (2007). Psychological foundations of trust. *Current Directions in Psychological Science*, *16*, 264–268. doi:10.1111/j.1467-8721.2007.00517.x
- Snyder, M., & Stukas, A. A. (1999). Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual Review of Psychology*, *50*, 273–303. doi:10.1146/annurev.psych.50.1.273
- Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, *35*, 656–666. doi:10.1037/0022-3514.35.9.656
- Tabak, J. A., & Zayas, V. (2012). The roles of featural and configural face processing in snap judgments of sexual orientation. *PLoS One*, *7*. doi:10.1371/journal.pone.0036671
- Taylor, A. B., MacKinnon, D. P., & Tein, J.-Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, *11*, 241–269. doi:10.1177/1094428107300344
- Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, *4*, 285–293. doi:10.1207/s15327965pli0104_1
- Todorov, A., & Porter, J. M. (2014). Misleading first impressions: Different for different facial images of the same person. *Psychological Science*, *25*, 1404–1417. doi:10.1177/0956797614532474
- Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*, 281–300. doi:10.1037/a0017908
- Vogt, D. S., & Colvin, R. C. (2003). Interpersonal orientation and the accuracy of personality judgments. *Journal of Personality*, *71*, 267–295. doi:10.1111/1467-6494.7102005
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x
- Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*, 1325–1331. doi:10.1177/0956797615590992
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151–175. doi:10.1037/0003-066X.35.2.151
- Zayas, V., Shoda, Y., & Ayduk, O. N. (2002). Personality in context: An interpersonal systems perspective. *Journal of Personality*, *70*, 851–900. doi:10.1111/1467-6494.05026
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, *12*, 399–413. doi:10.1037/1082-989X.12.4.399

Author Biographies

Gul Gunaydin is an assistant professor of Psychology at Bilkent University, Turkey. She received her PhD degree in Psychology from Cornell University. Her research focuses on how others from unknown individuals to romantic partners are mentally represented and the consequences of these representations for impression formation, interpersonal behavior, and emotion regulation. Email: gulgunaydin@bilkent.edu.tr cc; gg294@cornell.edu

Emre Selcuk is an assistant professor of Psychology at Middle East Technical University, Turkey. He received his PhD degree in Human Development from Cornell University. His research examines the formation, maintenance, and functions of attachment relationships across the life span, and the individual difference factors (e.g., perceived partner responsiveness) moderating these processes. Email: semre@metu.edu.tr

Vivian Zayas is an associate professor of Psychology and the director of the Personality, Attachment, and Control Laboratory at Cornell University, US. Her research examines the cognitive and affective processes that regulate behaviors within close relationships from a multilevel, interdisciplinary perspective. Email: vz29@cornell.edu

Handling Editor: Nickola Overall