Published in final edited form as:

Memory. 2015 October ; 23(7): 1039–1055. doi:10.1080/09658211.2014.951364.

Influences of Appearance-Behavior Congruity on Memory and Social Judgments

Brittany S. Cassidy and **Angela H. Gutchess**

Brandeis University

Abstract

Prior work shows that appearance-behavior congruity impacts memory and evaluations. Building upon prior work, we assessed influences of appearance-behavior congruity on source memory and judgment strength to illustrate ways congruity effects permeate social cognition. We paired faces varying on trustworthiness with valenced behaviors to create congruent and incongruent face-behavior pairs. Young and older adults remembered congruent pairs better than incongruent, but both were remembered better than pairs with faces rated average in appearance. This suggests that multiple, even conflicting, valenced cues improve memory over receiving fewer cues. Consistent with our manipulation of facial trustworthiness, congruity effects were present in the strength of trustworthiness-related but not dominance judgments. Subtle age differences emerged in congruity effects when learning about others, with older adults showing effects for approach judgments given both high and low arousal behaviors. Young adults had congruity effects for approach, prosociality, and trustworthiness judgments, given high arousal behaviors only. These findings deepen our understanding of how appearance-behavior congruity impacts memory for and evaluations of others.

Keywords

trustworthiness; impression formation; source memory; social judgments; aging

Impressions guide countless social decisions, and have inspired decades of research. This research shows people tend to agree on both appearance-based (e.g., why I interpret faces as conveying different traits) (Willis & Todorov, 2006; Zebrowitz & Montepare, 2008) and behavior-based (e.g., why I think someone is smart because she earned an A) (Todorov & Uleman, 2002, 2003) impressions, two distinct areas of work. The integration of appearance and behavioral cues is less studied, yet important because impression formation involves cue integration. Different cues are not always congruent, or consistent, meaning people's memories and actions could differ based on appearance-behavior congruity.

Appearance-behavior congruity is intriguing because people use both appearances (Todorov, 2008) and behaviors (McCarthy & Skowronski, 2011) to predict future behaviors and inform interactions. This appearance-behavior interplay affects many aspects of our lives. A

person's facial appearance (e.g., looking trustworthy or not), for instance, impacts the interpretation of ambiguous behavioral information (Hassin & Trope, 2000). As another example, people are more likely to agree with the positions of babyfaced over mature-faced individuals who have questionable trustworthiness after hearing their persuasive arguments, potentially because child-like faces convey honesty (Brownlow, 1992).

As previously studied (e.g., Nash, Bryer, & Schlaghecken, 2010), manipulating the valence conveyed by actor behaviors and faces varies appearance-behavior congruity. People associate subjectively trustworthy facial characteristics with positive traits, and untrustworthy features with negative (Oosterhof & Todorov, 2008), with facial width being one cue associated with trustworthiness perceptions (Stirrat & Perrett, 2010). People also infer traits from behaviors (Uleman, Saribay, & Gonzalez, 2008), associating them with their actors (Todorov & Uleman, 2002). Many impression formation experiments (e.g., Cassidy & Gutchess, 2012) control for facial characteristics across conditions, potentially downplaying how appearance-behavior congruity might influence interpreting and remembering others.

Recent work (Cassidy, Zebrowitz, & Gutchess, 2012; Kleider, Cavrak, & Knuycky, 2012; Nash et al., 2010; Rule, Slepian, & Ambady, 2012; Suzuki & Suga, 2010) explored appearance-behavior congruity effects in memory. Creating congruent (e.g., trustworthy face/positive behavior) and incongruent (e.g., trustworthy face/negative behavior) face-behavior pairs can lead to memory biases for both congruent (e.g., Nash et al., 2010) and incongruent (e.g., Suzuki & Suga, 2010) information. Conflicting findings correspond with outcome dependency hypothesis (Erber & Fiske, 1984; Stangor & McMillan, 1992): when perceiver outcomes are actor-dependent, incongruent information becomes privileged because perceivers want a better sense of prediction. Without actor-dependent outcomes, congruent information becomes privileged given motivation to maintain expectations. In memory, people might remember more incongruent person information in situations dependent on others (e.g., economic games). Without this dependency, people might better remember congruent information. Our studies explore congruity effects emerging from the latter.

Appearance-based inferences could function like schematic knowledge (i.e., associated information based on past experiences (Alba & Hasher, 1983)), contributing to congruity effects. Congruity effects may be akin to having better memory for schema-consistent sources when schemas act as decision cues (Bayen, Nakamura, Dupuis, & Yang, 2000). However, appearance-based information at test could cue *contextual* misattribution (e.g., this person performed a positive versus negative behavior), opposed to remembering *what* was presented (e.g., I saw this person before) (Johnson, Hashtroudi, & Lindsay, 1993). Thus, facial trustworthiness may be easier to use for decisions than remembering appearance-incongruent behaviors.

The present studies elaborated on prior work assessing congruity effects. First, we compared memory for congruent and incongruent appearance-behavior information to behaviors performed by comparatively average-faced actors (i.e., rated as not discernably trustworthy or untrustworthy). This extends previous work by investigating whether having two

enhanced diagnostic cues (i.e., both faces and behaviors) lead to better or worse memory for behaviors relative to receiving one cue (i.e., behavior only). Although people glean information from average faces (Zebrowitz & Montepare, 2008), faces perceived positively structurally resemble happiness, while those perceived negatively resemble anger (Said, Sebe, & Todorov, 2009). Average faces may have comparatively less structural resemblance to emotional expressions, giving fewer cues to overgeneralize traits. How memories formed from multiple conflicting cues compare to those from fewer is unclear.

Although past work manipulating appearance-behavior congruity has focused on memory (e.g., Nash et al., 2010), *when* effects emerge is unclear. Congruity effects could emerge when reconstructing information from memory. Effects could also manifest immediately after encountering congruent or incongruent information. Extant research supports the latter possibility (e.g., Hassin & Trope, 2000; Olivola & Todorov, 2010b). For instance, evaluating new faces as positive or negative depends on their similarity to ones previously paired with valenced information (Verosky & Todorov, 2010, 2013). Even when congruity does not impact actual ratings, it affects reaction time, with faster ratings toward congruent over incongruent face-trait pairs under time constraints (Rudoy & Paller, 2009). This suggests congruency impacts the ease of making person judgments. Although people may make similar overall *choices* based on congruent and incongruent cues (i.e., the binary choice to approach or avoid), we predicted judgment *strength* would be tempered by incongruent cues. In addition, we investigated effects across several judgments and under varying levels of behavioral arousal. These manipulations provide insight into congruity's impact on social evaluation depending on context.

Further contributing to the literature, we examined these topics in both young and older adults. Young and older adults make similar overgeneralizations from faces (Boshyan, Zebrowitz, Franklin, McCormick, & Carre, 2013; Zebrowitz, Franklin, Hillman, & Boc, 2013) and react similarly to structural resemblance to emotions (Franklin & Zebrowitz, 2013). Whether appearance-behavior congruity differentially impacts older adults is unknown. Beyond providing insight into the effect's generalizability across age groups, studying congruity effects with age is important given increased age-related fraud vulnerability (Ruffman, Murray, Halberstadt, & Vater, 2012). Age-related changes in interpreting deceptive person information may extend to considering potentially deceptive incongruent versus congruent appearance-behavior cues.

Aging could increase sensitivity to congruity. While young adults rely on schemas when making decisions (Hicks & Cockman, 2003; Marsh, Cook, & Hicks, 2006), older adults may rely on schemas to a greater extent (Hess & Follett, 1994; Mather & Johnson, 2003; Shi, Tang, & Liu, 2012). Older adults also activate more stereotyped information than young to explain social situations (Radvansky, Lynchard, & von Hippel, 2009) and process incongruent behavioral information less (Hess & Tate, 1991). Given facial stereotypes, older adults may reflect on behaviors less than young, exacerbating congruity effects.

However, aging could also *decrease* sensitivity to congruity. Older adults emphasize trait-diagnostic information in judgments of others more than young (Hess & Auman, 2001) and have increased sensitivity to behavioral information that might moderate first impressions

(Hess, Osowski, & Leclerc, 2005). This suggests greater social expertise with age (Hess, 2006). Attenuated congruity effects could emerge if older adults' inferences reflect behaviors better than young. We test these possibilities in the following studies.

Experiment 1

Although faces were *expressively* neutral in prior source memory work involving appearance-behavior congruity (e.g., Cassidy et al., 2012), subjectively trustworthy or untrustworthy faces may provide more diagnostic information (Said et al., 2009) than faces falling between these categories.

Two possibilities emerged for how memory for congruent and incongruent pairs would compare to average-face pairs. First, while congruent face-behavior pairs could be remembered over incongruent and average-face pairs, average-face pairs could be better remembered over incongruent because facial cues would not conflict with behaviors. Appearances might only be diagnostic to the extent that they do not conflict with behaviors. Second, congruent and incongruent pairs could be more accurately identified over average-face pairs if receiving enhanced diagnostic information from two sources versus one improves memory overall.

We also assessed age-related changes in the contribution of appearance-behavior congruity to source memory. We expected exacerbated congruity effects among older adults versus young, given age differences in schema reliance.

Method

Participants

Twenty-eight older (65–95 years old, 15 females; $M = 76.86$, $SD = 7.84$) and 28 young adults (17–21 years old, 18 females; $M = 18.57$, $SD = 1.00$) recruited from Brandeis University and the surrounding community participated. The 17-year-old received parental permission to participate. Older adults had Mini-Mental State Examination (MMSE) scores > 26 (Folstein, Folstein, & McHugh, 1975) ($M = 29.04$, $SD = 1.00$). Young ($M = 12.46$, $SD = 0.88$) and older ($M = 15.93$, $SD = 2.60$) adults differed in years of education, $t(54) = 6.69$, $p < 0.001$. Older adults ($M = 35.54$, $SD = 3.29$) had higher vocabulary scores (Shipley, 1986) than young ($M = 33.07$, $SD = 2.98$), $t(54) = 2.94$, $p = 0.005$. Young adults had faster processing speed ($M = 80.29$, $SD = 10.99$) than older adults ($M = 59.11$, $SD = 8.93$), $t(54) = 7.91$, $p < 0.001$, as measured by digit-comparison (Hedden et al., 2002), and higher letter-number sequencing scores (Wechsler, 1997) (young: $M = 11.25$, $SD = 1.94$; older adults: $M = 9.79$, $SD = 2.35$; $t(54) = 2.55$, $p = 0.01$).

Stimuli

Faces and behaviors—Pilot data from young and older adults was used to assign 48 male faces (24 young/24 older adults) drawn from the PAL database (Minear & Park, 2004) and 48 behaviors (24 positive/24 negative) drawn from a larger dataset (Somerville, Wig, Whalen, & Kelley, 2006) to conditions. Faces of each age group were evenly categorized as trustworthy, average, and untrustworthy. The trustworthy and untrustworthy faces did not

differ from average faces in distinctiveness. Analyses confirmed differences between the face categories and the valenced behaviors, and matched for other features (behavioral arousal and valence extremity). See supplemental materials for full details.

Face-behavior pairs—The 16 trustworthy, 16 average, and 16 untrustworthy faces were paired evenly with positive and negative behaviors (equally split among young and older faces in each trustworthiness category). Trustworthy face/positive behavior pairs and untrustworthy face/negative behavior pairs were categorized as “congruent.” Trustworthy face/negative behavior pairs and untrustworthy face/positive behavior pairs were categorized as “incongruent.” Average face/positive behavior pairs and average face/negative behavior pairs were categorized as “average-face” pairs. Two task versions counterbalanced the pairs. Faces paired with positive behaviors in one version were paired with negative in the other. Note that 16 face-behavior pairs per condition is consistent with prior related work (Nash et al., 2010) and meets the recommendation for the minimum number of stimuli per condition in counterbalanced designs (Westfall, Kenny, & Judd, in press).

Procedure

Encoding—Unaware of a memory task, participants were told they were participating in a task assessing reaction times and social cognition. Stimuli were presented via E-Prime (Psychology Software Tools, Pittsburgh, PA). Participants viewed each face alone for two seconds. Participants pressed “1” after viewing the face. Then, they saw the same face paired with a behavior for five seconds. Participants pressed “2” after reading the sentence (Figure 1a). Participants were told their responses would not advance stimuli, but that they should respond quickly. Responses ensured attention to stimuli.

There were four blocks of twelve trials each in each run, with same-age faces in each block and six seconds between each block. Block order was counterbalanced among the task versions (i.e., half of participants saw older faces in the first block). To improve performance, pairs appeared twice, once per run in a random order. After encoding, participants completed the digit comparison measure (Hedden et al., 2002).

Retrieval—Participants completed a self-paced task. Participants viewed all faces from encoding, with no new faces, one at a time in a random order. Participants were told they would view faces from the previous task, with no new ones. The words “Positive” and “Negative” appeared below each face. Participants indicated their behavior-based impressions as positive or negative (Figure 1b). Memory was scored based on prior behavior categorizations as positive or negative, similar to other work (Cassidy et al., 2012; Nash et al., 2010).

Unbiased hit rate

We used unbiased hit rates, proposed by Wagner (1993) and previously employed in assessing appearance-behavior congruity influences on memory (Suzuki & Suga, 2010) to control for response biases (e.g., a tendency to respond “positive” for trustworthy faces). We calculated unbiased hit rates by multiplying hit rates by response bias terms. For example, the hit rate for congruent pairs would be the proportion of the number of faces (collapsing

across trustworthy and untrustworthy) previously associated with congruent behaviors (collapsing across positive and negative) that were correctly classified to the total number of congruent pairs in the study. The response bias term would be the proportion of the number of congruent pairs correctly classified to the total number of faces classified congruently with appearance. We calculated unbiased hit rates for congruent, incongruent, and average-face pairs.

Results

Memory

We analyzed memory in a 2 (Age Group: Young, Old) $\times$ 3 (Congruity: Congruent, Incongruent, Average-face) mixed ANOVA. All participants performed above chance. A congruity effect existed, $F(2, 108) = 8.57, p < 0.001, \eta_p^2 = 0.14$. Congruent pairs ($M = 0.54, SD = 0.15$) were better remembered than incongruent ($M = 0.51, SD = 0.17$), $F(1, 54) = 4.58, p = 0.04, \eta_p^2 = 0.08$ and average-face pairs ($M = 0.46, SD = 0.18$), $F(1, 54) = 13.82, p < 0.001, \eta_p^2 = 0.20$. Incongruent pairs were remembered better than average-face ones, $F(1, 54) = 5.21, p = 0.03, \eta_p^2 = 0.09$. Contrary to prediction, no Age Group $\times$ Congruity interaction emerged, $F(2, 108) = 0.03, p = 0.97, \eta_p^2 < 0.01$. No main effect of Age Group existed, $F(1, 54) = 0.08, p = 0.78, \eta_p^2 < 0.01$. See Table 1 for full descriptive statistics for each age group. Note that power analyses conducted in G*Power (Mayr, Erdfelder, Buchner, & Faul, 2007) for sample size estimation suggested a total sample of 34 to detect effects using $\alpha = 0.05$, power = 0.80, and Cohen's $f = 0.23$ ($\eta_p^2 = 0.05$).

Exploring the contributions of behavioral valence and facial trustworthiness to memory

Our primary purpose was to examine effects appearance-behavior congruity on source memory and related effects of aging. Other work illustrates age-related changes in processing of valenced material. Our primary analysis would not capture this possibility. While young may focus on incoming negative cues, older adults may attend to more positive information (Mather & Carstensen, 2005; Spaniol, Voss, & Grady, 2008). We tested this idea in a 2 (Age Group: Young, Old) $\times$ 3 (Facial Trustworthiness: Trustworthy, Untrustworthy, Average) $\times$ 2 (Behavioral Valence: Positive, Negative) mixed ANOVA.

Facial Trustworthiness impacted memory, $F(2, 108) = 5.95, p = 0.004, \eta_p^2 = 0.10$. Behaviors of average faces ($M = 0.46, SD = 0.10$) were remembered less well than those of trustworthy ($M = 0.52, SD = 0.10$), $F(1, 54) = 5.57, p = 0.02, \eta_p^2 = 0.09$, and untrustworthy faces ($M = 0.55, SD = 0.46$), $F(1, 54) = 10.56, p = 0.002, \eta_p^2 = 0.16$. Memory of trustworthy and untrustworthy faces' behaviors did not differ, $F(1, 54) = 1.64, p = 0.21, \eta_p^2 = 0.03$.

A marginal interaction existed between Facial Trustworthiness and Behavioral Valence, $F(2, 108) = 3.07, p = 0.05, \eta_p^2 = 0.05$. People remembered more positive ($M = 0.54, SD = 0.07$) than negative ($M = 0.50, SD = 0.08$) behaviors from trustworthy faces, $F(1, 54) = 5.92, p = 0.02, \eta_p^2 = 0.10$. Memory did not differ between valenced behaviors of untrustworthy (Positive: $M = 0.55, SD = 0.08$; Negative $M = 0.56, SD = 0.08$; $F(1, 54) = 0.78, p = 0.38, \eta_p^2 = 0.01$), or average (Positive: $M = 0.47, SD = 0.07$; Negative $M = 0.45, SD = 0.07$; $F(1, 54) = 1.55, p = 0.22, \eta_p^2 = 0.03$) faces. No other effects were significant, $ps > 0.12$.

Discussion

Congruent pairs were remembered over incongruent and average-face pairs, evidencing a congruity effect in memory and extending prior work. This demonstrates the extent of memory enhancements or impairments given appearance-behavior congruity. Even when inconsistent, receiving multiple streams of enhanced valenced information improves memory over receiving fewer. If both faces and behaviors convey clearly valenced diagnostic information, they may be better attended to, leading to richer person representations and more accurate source memory decisions.

Interestingly, our findings may connect to more general memory work on dual coding, in which two different streams of descriptive information (e.g., verbal and visual) lead to enhanced memory over fewer streams (for a review, see Paivio, 1991). Previous work has also found better facial recognition with increased task emotionality (Mueller, Heesacker, Ross, & Nicodemus, 1983). Our task may be similar in that the social value of stimuli could be increased based on the diagnostic value of appearance- and behavior-derived traits. Alternatively, our findings may reflect a dilution effect (Nisbett, Zukier, & Lemley, 1981). Comparably less trait-diagnostic information from appearances (i.e., appearing average compared to more extremely trustworthy or untrustworthy) could have weakened the predictive value of valenced behaviors, reducing memory compared to when both faces and behaviors conveyed richer trait-diagnostic information. Average faces might also contain more ambiguous cues than overtly trustworthy or untrustworthy ones, weakening their diagnostic value and reducing memory. Future work could test these possibilities.

Contrary to our prediction, age-invariant memory performance suggests older adults remember incongruent appearance-behavior information like young. This is consistent with work showing young and older adults similarly remember socioemotional information (Cassidy & Gutchess, 2012; May, Rahhal, Berry, & Leighton, 2005; Rahhal, May, & Hasher, 2002; Todorov & Olson, 2008), in contrast to age-related declines in non-social memory tasks (review by Grady & Craik, 2000). However, it is also possible that the self-paced retrieval task may have allowed time for older adults to resolve incongruity. A constrained task could elicit age differences, as increasing cognitive load reduces the ability to resolve incongruent information (Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999).

Finally, remembering congruent over incongruent *trustworthy* face-behavior pairs drove the present congruity effect. Untrustworthy faces are prioritized in memory, potentially reflecting an evolved “cheater detection system” (Mealey, Daoood, & Krage, 1996). Indeed, people have better memory for untrustworthy over trustworthy faces (Mueller, Thompson, & Vogel, 1988), regardless of whether the faces are encoded with behaviors (Rule et al., 2012). Negative facial cues could warrant the allocation of attention to associated behaviors, increasing memory. When encountering positive facial cues, schema-consistent expectations might lead to the prioritization of positive behaviors over negative. Additionally, these findings complement recent work showing that both trustworthy and untrustworthy faces are better remembered than trustworthy-neutral ones (Mattarozzi, Todorov, & Codispoti, 2014).

Experiment 2

Like prior work, Experiment 1 assessed contributions of appearance-behavior congruity to memory. Other work (e.g., Hassin & Trope, 2000) suggests appearance-behavior congruity effects are not solely memory phenomena; thus, we investigated whether congruity impacts judgment strength immediately after learning about others, without relying on memory. We predicted judgments to be stronger toward congruent over incongruent face-behavior pairs, due to the convergence of multiple streams of information. This prediction could impact important social outcomes. For instance, the strength of initial reactions could impact whether someone receives a second job interview.

Like Experiment 1, we expected age-related exacerbation of congruity effects given increased schema reliance with age. However, given other work on aging and impression formation (Hess et al., 2005), these effects could be attenuated with age, should older adults be particularly sensitive to incoming behavioral information.

As a secondary goal, we examined two factors potentially modulating how congruity impacts judgment strength. First, we tested how effects might be constrained by the *type* of judgment at hand. To test this, we assessed the presence of effects for trustworthiness-related judgments for social interactions involving direct interaction (i.e., approach behavior), indirect interactions (i.e., prosociality behavior, in which one donates to an individual but does not have to meet that person), and personality judgments (i.e., perceived trustworthiness). We used a dominance judgment as a control. Trustworthiness and dominance fall on separable dimensions of face evaluation (Oosterhof & Todorov, 2008). Because we manipulated facial trustworthiness and not dominance, we predicted congruity effects would hold for the three trustworthiness-related judgments but not the dominance judgment.

Second, we assessed how *behavioral arousal* impacts congruity effects. Arousal influences social cognition. For instance, events eliciting heightened arousal (i.e., an increased emotional response) are often better remembered (Cahill & McGaugh, 1998), and arousal influences subsequent judgments (Peters, Vastfjall, Garling, & Slovic, 2006). Given arousing behaviors, we predicted similar judgments for congruent and incongruent face-behavior pairs, reflecting behavioral emphasis. Given low arousal behaviors (i.e., eliciting little emotional response), we predicted stronger judgments toward congruent over incongruent pairs. Here, appearances might be more likely to distort judgments.

Method

Participants

Thirty-two older (63–93 years old, 23 females; $M = 75.38$, $SD = 7.94$) and 32 young adults (18–22 years old, 20 females; $M = 18.38$, $SD = 0.91$) recruited from Brandeis University and the surrounding community participated. Older adults had MMSE scores > 26 (Folstein et al., 1975) ($M = 29.22$, $SD = 1.01$). Older adults ($M = 16.67$, $SD = 2.22$) had more years of education than young ($M = 13.08$, $SD = 1.17$), $t(62) = 8.13$, $p < 0.001$. Older adults ($M = 36.13$, $SD = 4.35$) had higher vocabulary scores (Shipley, 1986) than young ($M = 30.47$, SD

= 4.09), $t(62) = 5.36, p < 0.001$. Young had faster processing speed ($M = 81.47, SD = 14.67$) than older adults ($M = 60.69, SD = 11.91$), $t(62) = 6.22, p < 0.001$ (Hedden et al., 2002), and higher letter-number sequencing scores (Wechsler, 1997) (young: $M = 11.97, SD = 2.38$; older adults: $M = 9.72, SD = 2.80$; $t(62) = 3.47, p = 0.001$).

Stimuli

Faces and behaviors—Pilot data from young and older adults was used to assign 80 young male faces (evenly split by trustworthy/untrustworthy) generated via FaceGen version 3.1 using a model developed by Oosterhof and Todorov (2008) and 80 behaviors (evenly split by positive/negative valence and high/low arousal) from previous aging work (Cassidy & Gutchess, 2012) to conditions. Analyses confirmed differences between trustworthy and untrustworthy faces, and the positive and negative behaviors varying on arousal. Analyses also assessed differences in stimuli on other relevant features (perceived face age, perceived face valence, behavioral valence, behavioral valence extremity), with no differences impacting the present study. See supplemental materials for full details.

Face-behavior pairs—The 40 trustworthy and 40 untrustworthy faces were evenly paired with congruent (trustworthy-positive and untrustworthy-negative) or incongruent (trustworthy-negative or untrustworthy-positive) high and low arousal behaviors. Faces were paired with high or low arousal congruent or incongruent behaviors across four task versions.

Procedure

Participants were told this was a self-paced task about how people get to know others. Stimuli were presented via E-Prime (Psychology Software Tools, Pittsburgh, PA). Participants viewed each face alone for two seconds. Then, participants saw the same face paired with a behavior for five seconds (Figure 2). After the five-second interval, participants made four self-paced judgments: Approach (“Would I approach this person?”), Prosociality (“If in need, how much would I donate to this person?”), Trustworthiness (“How trustworthy is this person?”) and Dominance (“How dominant is this person?”). Responses were entered via 9-point scales. Question order varied by version (i.e., each question was asked first, second, third, or fourth). There were eight blocks of ten trials each, with six seconds between each block.

Judgment Strength

To compare judgments made about different types of congruent and incongruent pairs (e.g., collapsing across trustworthy face/positive behavior and untrustworthy face/negative behavior), we calculated judgment strength separately for the four judgments. Specific responses were subtracted from the mean of each participant’s responses for each judgment type. To compare judgment strength, we took the absolute values of these subtractions. This accounts for individual differences in scale use, and is similar to calculations from aging work on trait attribution (Follett & Hess, 2002).

Results

We analyzed judgment strength in a 2 (Age Group: Young, Old) $\times$ 4 (Judgment: Approach, Prosociality, Trustworthiness, Dominance) $\times$ 2 (Congruity: Congruent, Incongruent) $\times$ 2 (Behavioral Arousal: High, Low) mixed ANOVA. For brevity, findings unrelated to congruity have been relegated to footnotes¹. Note that power analyses conducted in G*Power suggested a total sample size of 40 to detect effects using $\alpha = 0.05$, power = 0.80, and Cohen's $f = 0.23$ ($\eta_p^2 = 0.05$).

Influence of congruity on judgment strength

Judgments were stronger toward congruent ($M = 1.95$, $SD = 0.53$) over incongruent ($M = 1.85$, $SD = 0.50$) pairs, $F(1, 62) = 13.79$, $p < 0.001$, $\eta_p^2 = 0.18$. The anticipated interaction between Judgment and Congruity emerged, $F(3, 186) = 4.67$, $p = 0.004$, $\eta_p^2 = 0.07$ (Figure 3). Congruent pairs elicited stronger judgments than incongruent for approach (Congruent: $M = 2.02$, $SD = 0.66$; Incongruent: $M = 1.85$, $SD = 0.62$; $F(1, 62) = 16.87$, $p < 0.001$, $\eta_p^2 = 0.21$), prosociality (Congruent: $M = 1.90$, $SD = 0.68$; Incongruent: $M = 1.81$, $SD = 0.67$; $F(1, 62) = 5.15$, $p = 0.03$, $\eta_p^2 = 0.08$), and trustworthiness (Congruent: $M = 2.01$, $SD = 0.55$; Incongruent: $M = 1.90$, $SD = 0.53$; $F(1, 62) = 10.30$, $p = 0.002$, $\eta_p^2 = 0.14$) judgments. Dominance judgments were unaffected by congruency (Congruent: $M = 1.86$, $SD = 0.56$; Incongruent: $M = 1.86$, $SD = 0.55$; $F(1, 62) = 0.01$, $p = 0.92$, $\eta_p^2 < 0.01$).

Influence of congruity qualified by behavioral arousal

Behavioral Arousal qualified the relationship between Judgment and Congruity, $F(3, 186) = 3.19$, $p = 0.03$, $\eta_p^2 = 0.05$. We report findings for each judgment.

Approach—Judgments toward high arousal-congruent pairs ($M = 2.31$, $SD = 0.73$) were stronger than for incongruent pairs ($M = 2.08$, $SD = 0.72$), $F(1, 62) = 15.61$, $p < 0.001$, $\eta_p^2 = 0.20$. This also occurred among low arousal pairs ($M_{congruent} = 1.71$, $SD_{congruent} = 0.63$; $M_{incongruent} = 1.61$, $SD_{incongruent} = 0.62$), $F(1, 62) = 5.31$, $p = 0.03$, $\eta_p^2 = 0.08$.

Prosociality—Judgments toward high arousal-congruent pairs ($M = 2.20$, $SD = 0.76$) were stronger than incongruent ($M = 2.04$, $SD = 0.75$), $F(1, 62) = 8.21$, $p = 0.01$, $\eta_p^2 = 0.12$. Low arousal congruent ($M = 1.60$, $SD = 0.66$) and incongruent ($M = 1.58$, $SD = 0.66$) pairs did not differ, $F(1, 62) = 0.14$, $p = 0.71$, $\eta_p^2 < 0.01$.

¹Judgments were stronger given high ($M = 2.14$, $SD = 0.54$) versus low arousal ($M = 1.66$, $SD = 0.50$), $F(1, 62) = 214.83$, $p < 0.001$, $\eta_p^2 = 0.78$, and qualified by Judgment, $F(3, 186) = 25.86$, $p < 0.001$, $\eta_p^2 = 0.29$. Given high arousal, dominance judgments ($M = 1.98$, $SD = 0.57$) were weaker than both approach ($M = 2.20$, $SD = 0.69$; $F(1, 62) = 11.70$, $p = 0.001$, $\eta_p^2 = 0.16$) and trustworthiness ($M = 2.26$, $SD = 0.58$; $F(1, 62) = 19.76$, $p < 0.001$, $\eta_p^2 = 0.24$) judgments. Trustworthiness judgments were marginally stronger than prosociality ($M = 2.12$, $SD = 0.72$) judgments, $F(1, 62) = 2.95$, $p = 0.09$, $\eta_p^2 = 0.05$. There was no difference between dominance and prosociality judgments ($F(1, 62) = 2.80$, $p = 0.10$, $\eta_p^2 = 0.04$), and between approach and trustworthiness ($F(1, 62) = 1.11$, $p = 0.30$, $\eta_p^2 = 0.02$) or prosociality ($F(1, 62) = 1.38$, $p = 0.25$, $\eta_p^2 = 0.02$) judgments. Given low arousal, participants made stronger dominance ($M = 1.74$, $SD = 0.54$) than prosociality ($M = 1.59$, $SD = 0.64$; $F(1, 62) = 4.35$, $p = 0.04$, $\eta_p^2 = 0.07$) judgments. Participants made marginally stronger dominance than trustworthiness judgments ($M = 1.65$, $SD = 0.53$), $F(1, 62) = 2.88$, $p = 0.095$, $\eta_p^2 = 0.04$. There was no difference between dominance and approach ($M = 1.66$, $SD = 0.60$) judgments, $F(1, 62) = 1.69$, $p = 0.20$, $\eta_p^2 = 0.03$. There were no differences between trustworthiness judgments and approach ($F(1, 62) = 0.12$, $p = 0.73$, $\eta_p^2 < 0.01$) and prosociality ($F(1, 62) = 0.93$, $p = 0.34$, $\eta_p^2 = 0.02$) judgments, and between approach and prosociality judgments ($F(1, 62) = 1.36$, $p = 0.25$, $\eta_p^2 = 0.02$). No other findings in the experiment were significant, $ps > 0.14$.

Trustworthiness—Judgments toward high arousal-congruent pairs ($M = 2.34$, $SD = 0.63$) were stronger than incongruent ($M = 2.04$, $SD = 0.58$), $F(1, 62) = 9.98$, $p = 0.002$, $\eta_p^2 = 0.14$. Low arousal-congruent ($M = 1.67$, $SD = 0.58$) and incongruent ($M = 1.63$, $SD = 0.55$) pairs did not differ, $F(1, 62) = 0.90$, $p = 0.35$, $\eta_p^2 < 0.01$.

Dominance—High arousal-congruent ($M = 1.97$, $SD = 0.62$) and incongruent ($M = 2.00$, $SD = 0.61$) pairs did not differ, $F(1, 62) = 0.37$, $p = 0.55$, $\eta_p^2 < 0.01$. Low arousal-congruent ($M = 1.76$, $SD = 0.60$) and incongruent ($M = 1.72$, $SD = 0.57$) pairs did not differ, $F(1, 62) = 0.56$, $p = 0.46$, $\eta_p^2 < 0.01$.

Age differences in the influence of congruity on judgment strength

The predicted Age Group $\times$ Congruity interaction was not significant, $F(1, 62) = 0.04$, $p = 0.85$, $\eta_p^2 < 0.01$. However, there was a marginal interaction between Age Group, Congruity, Judgment, and Arousal, $F(3, 186) = 2.57$, $p = 0.06$, $\eta_p^2 = 0.04$ (Figure 4). We consider findings for each judgment.

Approach—Both young ($M_{congruent} = 2.37$, $SD_{congruent} = 0.68$; $M_{incongruent} = 2.13$, $SD_{incongruent} = 0.68$) and older ($M_{congruent} = 2.27$, $SD_{congruent} = 0.79$; $M_{incongruent} = 2.04$, $SD_{incongruent} = 0.76$) adults had stronger judgments for high arousal-congruent over incongruent face-behavior pairs (Young: $F(1, 31) = 7.57$, $p = 0.01$, $\eta_p^2 = 0.20$; Old: $F(1, 31) = 8.12$, $p = 0.01$, $\eta_p^2 = 0.21$). Older adults had stronger judgments for low arousal-congruent ($M = 1.65$, $SD = 0.64$) over incongruent ($M = 1.48$, $SD = 0.59$) pairs, $F(1, 31) = 7.88$, $p = 0.01$, $\eta_p^2 = 0.20$. There was no difference between low arousal-congruent ($M = 1.79$, $SD = 0.62$) and incongruent ($M = 1.74$, $SD = 0.62$) among young, $F(1, 31) = 0.43$, $p = 0.52$, $\eta_p^2 = 0.01$.

Prosociality—Young adults had stronger judgments toward high arousal-congruent ($M = 2.26$, $SD = 0.77$) over incongruent ($M = 2.00$, $SD = 0.66$) pairs, $F(1, 31) = 10.91$, $p = 0.002$, $\eta_p^2 = 0.26$. In older adults, high arousal-congruent ($M = 2.13$, $SD = 0.75$) and incongruent ($M = 2.08$, $SD = 0.62$) did not differ, $F(1, 31) = 0.51$, $p = 0.48$, $\eta_p^2 = 0.02$. Low arousal-congruent and incongruent behaviors did not differ for young ($M_{congruent} = 1.60$, $SD_{congruent} = 0.65$; $M_{incongruent} = 1.59$, $SD_{incongruent} = 0.66$; $F(1, 31) = 0.22$, $p = 0.64$, $\eta_p^2 < 0.01$) or older adults ($M_{congruent} = 1.57$, $SD_{congruent} = 0.57$; $M_{incongruent} = 1.56$, $SD_{incongruent} = 0.66$; $F(1, 31) < 0.01$, $p = 0.99$, $\eta_p^2 < 0.01$).

Trustworthiness—Young adults had stronger judgments toward high arousal-congruent ($M = 2.32$, $SD = 0.56$) over incongruent ($M = 2.10$, $SD = 0.55$) pairs, $F(1, 31) = 9.72$, $p = 0.01$, $\eta_p^2 = 0.24$. In older adults, high arousal-congruent ($M = 2.35$, $SD = 0.76$) and incongruent ($M = 2.36$, $SD = 0.63$) did not differ, $F(1, 31) = 1.76$, $p = 0.19$, $\eta_p^2 = 0.05$. Low arousal-congruent and incongruent behaviors did not differ for young ($M_{congruent} = 1.68$, $SD_{congruent} = 0.57$; $M_{incongruent} = 1.66$, $SD_{incongruent} = 0.55$; $F(1, 31) = 0.06$, $p = 0.81$, $\eta_p^2 < 0.01$) or older adults ($M_{congruent} = 1.67$, $SD_{congruent} = 0.57$; $M_{incongruent} = 1.60$, $SD_{incongruent} = 0.55$; $F(1, 31) = 2.09$, $p = 0.16$, $\eta_p^2 = 0.06$).

Dominance—Judgments of high arousal-congruent and incongruent pairs did not differ for young ($M_{congruent} = 2.00$, $SD_{congruent} = 0.61$; $M_{incongruent} = 2.11$, $SD_{incongruent} = 0.61$; $F(1, 31) = 1.83$, $p = 0.19$, $\eta_p^2 = 0.06$) or older adults ($M_{congruent} = 1.93$, $SD_{congruent} = 0.62$; $M_{incongruent} = 1.88$, $SD_{incongruent} = 0.61$; $F(1, 31) = 0.50$, $p = 0.49$, $\eta_p^2 = 0.02$). Low arousal-congruent and incongruent did not differ for young ($M_{congruent} = 1.83$, $SD_{congruent} = 0.60$; $M_{incongruent} = 1.78$, $SD_{incongruent} = 0.57$; $F(1, 31) = 0.45$, $p = 0.51$, $\eta_p^2 = 0.01$) or older adults ($M_{congruent} = 1.68$, $SD_{congruent} = 0.60$; $M_{incongruent} = 1.66$, $SD_{incongruent} = 0.57$; $F(1, 31) = 0.14$, $p = 0.72$, $\eta_p^2 < 0.01$).

Exploring the contributions of behavioral valence and facial trustworthiness to judgment strength

Like Experiment 1, we assessed potential contributions of behavioral valence and facial trustworthiness to judgment strength. Negative information may be more salient than positive (Rozin & Royzman, 2001). If negativity biases attenuate with age (e.g., Wood & Kisley, 2006), older adults' judgments may be similar regardless of behavioral valence. We tested this in a 2 (Age Group: Young, Old) $\times$ 2 (Facial Trustworthiness: Trustworthy, Untrustworthy) $\times$ 2 (Behavioral Valence: Positive, Negative) mixed ANOVA.

There were stronger judgments for trustworthy ($M = 1.94$, $SD = 0.52$) versus untrustworthy ($M = 1.86$, $SD = 0.54$) faces, $F(1, 62) = 4.54$, $p = 0.04$, $\eta_p^2 = 0.07$. Participants made stronger judgments given negative ($M = 1.81$, $SD = 0.52$) versus positive ($M = 1.99$, $SD = 0.51$) behaviors, $F(1, 62) = 44.98$, $p < 0.001$, $\eta_p^2 = 0.42$.

An interaction between Facial Trustworthiness and Behavioral Valence emerged, $F(1, 62) = 12.40$, $p = 0.001$, $\eta_p^2 = 0.17$. Congruent positive pairs ($M = 1.90$, $SD = 0.57$) elicited stronger judgments than pairs with positive behaviors but untrustworthy faces ($M = 1.73$, $SD = 0.56$), $F(1, 62) = 9.80$, $p = 0.003$, $\eta_p^2 = 0.44$. Congruent negative pairs ($M = 1.99$, $SD = 0.55$) and pairs with negative behaviors but trustworthy faces ($M = 1.98$, $SD = 0.51$), did not differ, $F(1, 62) = 0.08$, $p = 0.78$, $\eta_p^2 < 0.01$.

An Age Group by Behavioral Valence interaction existed, $F(1, 62) = 4.63$, $p = 0.04$, $\eta_p^2 = 0.07$. Young adults had stronger judgments given negative ($M = 2.05$, $SD = 0.51$) versus positive ($M = 1.82$, $SD = 0.53$) behaviors, $F(1, 31) = 57.35$, $p < 0.001$, $\eta_p^2 = 0.65$. Older adults also showed this pattern (Negative: $M = 1.92$, $SD = 0.51$; Positive $M = 1.80$, $SD = 0.53$), $F(1, 31) = 7.88$, $p = 0.01$, $\eta_p^2 = 0.20$. A larger effect size for young versus older adults means behavioral valence accounts for less variance in judgment strength with age. This suggests age-related attenuation in making stronger judgments when behaviors are negative versus positive. No other effects were significant, $ps > 0.55$.

Discussion

Judgments were stronger for congruent over incongruent face-behavior pairs, extending prior work by demonstrating congruity effects immediately after learning about others. This may inform findings of appearance-congruent behavioral outcomes. For instance, looking competent improves the odds of winning elections (Olivola & Todorov, 2010a). Performing appearance-congruent behaviors may magnify this effect.

Congruity impacted the strength of three trustworthiness-related judgments but not dominance judgments. This suggests congruity impacts passing personality judgments (“How trustworthy is this person?”) as well as judgments potentially evoking deeper processing given potential social interaction (“Would I approach this person?”). Congruity may matter most when decisions directly relate to facial features. Thus, trustworthiness may contribute to judgment strength most when decisions involve trustworthiness and valence. Future work could test this idea by comparing trustworthiness and competence judgment strength toward valenced face-behavior pairs. Competence may be distinct from trustworthiness/warmth (Fiske, Cuddy, & Glick, 2007), yet is more rigidly valenced than dominance because dominant *and* passive behaviors vary in valence. If congruity effects depend on manipulated facial features, manipulating trustworthiness should not result in effects on competence judgments.

Behavioral arousal qualified the relationship between congruity and judgment type differently than anticipated. We expected effects for pairs including low versus high arousal behaviors. Instead, we found effects in the three trustworthiness-related judgments for pairs with *high*, not low, arousal behaviors. An effect also existed among approach judgments given low arousal behaviors. This suggests high arousal appearance-congruent behaviors *enhance* judgment strength relative to incongruent ones. Learning appearance-congruent behaviors could enhance impressions, similar to how people tend to accept information confirming their opinions (Lord, Ross, & Lepper, 1979).

Fewer differences for low arousal behaviors may reflect less diagnostic aspects of low arousal information (Skowronski & Carlston, 1987). Alternatively, attention to relevant facial characteristics might also be enhanced by congruent high arousal behaviors, resulting in stronger judgments. Previous work is consistent with this idea. For instance, people have stronger evaluations towards those who look similar to people paired with valenced information (Verosky & Todorov, 2010, 2013).

Age-related influences on congruity effects differed from expectations. While both age groups had stronger approach judgments for high arousal-congruent over incongruent pairs, older adults had the same relationship for low arousal pairs. In contrast, young, but not older, adults had stronger prosociality and trustworthiness judgments for high arousal-congruent versus incongruent pairs. This suggests that for the above-described relationship between judgment, congruity, and behavioral arousal, older adults drove effects for approach judgments, and young adults drove effects elicited by prosociality and trustworthiness judgments.

A stronger effect for older adults in approach judgments complements work showing that despite fewer reported friendships (Chown, 1981), emotional closeness increases with age (Carstensen, 1992). If social partner selectivity increases with age, congruity effects in approach decisions could optimize the likelihood of successful interactions. Although unpredicted, appearance-behavior congruity may become more important for considering social interactions with age, and less important for decisions not involving the self. Future work could test this by manipulating the extent of personal implications from different judgments.

The presence of negative behaviors attenuated congruity effects. Negative information captures attention more than positive (Pratto & John, 1991; Skowronski & Carlston, 1989). Negative behaviors are also more diagnostic in the morality domain, with positive behaviors more diagnostic in the ability domain (Skowronski & Carlston, 1987). It would be worthwhile to assess whether effects weaken in the presence of negative morality-related and positive ability-related behaviors. This relationship also differs from the exploratory findings of Experiment 1. There, a congruity effect emerged given trustworthy faces, with similar memory for behaviors of untrustworthy faces. Here, an effect emerged for positive, but not negative, behaviors. While untrustworthy faces could be prioritized in memory, negative behaviors might contribute more to the strength of immediate judgments. Moreover, age-related attenuation complements work showing age-related shifts toward focusing on more positive material. Although less stark than the reversals in focus shown in emotion work (see Carstensen & Mikels, 2005; Mather & Carstensen, 2005), a tendency to judge negative behaviors more strongly than positive may attenuate with age.

General Discussion

The present experiments extend the literature on appearance-behavior congruity in three ways. First, Experiment 1 replicates work (e.g., Cassidy et al., 2012; Kleider et al., 2012; Nash et al., 2010) showing better memory for congruent over incongruent face-behavior pairs, while uniquely showing that congruent and incongruent pairs are better remembered than pairs with average faces. This suggests that even when appearance and behavioral cues conflict, they improve memory beyond when receiving less information. In Experiment 2, we used a similar manipulation to illustrate that congruity effects are not specific to memory. We show similar congruity effects immediately upon learning about others. Finally, we show congruity effects among older adults while revealing subtle age differences in their employment.

An important consideration of these studies is that faces were always shown before behaviors. If early-presented information has more impact on impressions than subsequently presented material (Anderson & Barrios, 1961), appearances anchored impressions in the present work. Comparing memory and judgment strength after manipulating face/behavior presentation order could test whether early-presented information has the most impact on impressions. If order matters, showing behaviors first may increase memories of them, reducing distortion from appearances.

Age-related influences on congruity effects were more complex than predicted. In Experiment 2, congruity influenced older adults across approach judgments, and young across trustworthiness-related judgments given high arousal behaviors. Older adults could be particularly sensitive to congruity in the context of social interactions. Connecting to prior work, older adults may indeed make better use of trait diagnostic information than young in some situations (Hess & Auman, 2001; Hess et al., 2005), while sometimes using schemas more than young (Mather & Johnson, 2003; Shi et al., 2012) to optimize the likelihood of positive social interactions.

A lack of overall age differences in the present studies may inform work evidencing differential processing of negative facial cues with age. When shown without additional cues (e.g., behaviors), older adults perceive untrustworthy faces more positively than young (Castle et al., 2012; Zebrowitz et al., 2013). Although this positive skew suggests difficulty in perceiving trustworthiness-based appearance-behavior congruity, our data do not reflect this. The processing of valenced cues from multiple sources may be preserved with age in that the presence of behaviors could induce closer consideration of faces, reducing age differences. However, in Experiment 2, we did find age-related attenuation for stronger judgments given negative versus positive behaviors. While age differences in processing negative cues may exist, the presence of behaviors could ameliorate differences in judgments about faces. However, choosing stimuli based on age-matched ratings, like in Experiment 1, could mask the detection of age differences. It would be worth testing for age differences in a naturalistic setting, where stimuli are uncontrolled, to examine this idea further.

Because inconsistent information can pop-out to perceivers, finding better memory and stronger judgments for congruent over incongruent information might seem counterintuitive. Indeed, not all studies evidence congruity effects (for a review, see Stangor & McMillan, 1992). We predicted congruity (rather than incongruity) effects in the present studies because actor behaviors did not directly impact perceivers. Other explanations, including schema-expectancy strength and relatively large number of stimuli, could also account for congruity effects, however. Individual differences also predict the size of effects, as those believing that traits are fixed remember more stereotype-congruent information (Plaks, Stroessner, Dweck, & Sherman, 2001). However, people learn others' traits better after exposure to congruent stimuli (Carlston & Skowronski, 1994), suggesting a strong role of congruity in impression formation. One reason why these findings might seem counterintuitive is that people often claim objectivity in decisions despite their actual behavior (Hansen, Gerbasi, Todorov, Kruse, & Pronin, 2014). Gaining a richer understanding of congruity effects is therefore important.

Like related work, these studies have implications for the legal system. Appearances contribute to jury behavior, as facial maturity impacts the odds of conviction for negligent or intentional crimes (Zebrowitz & McDonald, 1991). If a person *looks* like he would have committed a crime and is convicted, that person may be given a harsher sentence than someone who committed a similar crime but did not look the part. Moreover, it may be easier to remember negative information about a person who allegedly committed a crime than positive, but conflicting, information. This suggests congruity could affect juror and eyewitness impressions.

Taken together, these studies enhance our understanding of congruity effects in social cognition across age. As appearances impact many outcomes, gaining a greater understanding of how appearance-behavior congruity impacts cognition may more broadly inform our understanding of human interaction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Avi Aizenman, Rebekah LaFontant, Alina Liu, Trent Judis, and Tali Friedland for research assistance. We also thank Xiaodong Liu for his suggestions for statistical analyses. The National Institute on Aging grant R21 AG032382 (to A.H.G.) and an NSF graduate fellowship (to B.S.C.) supported this work.

References

- Alba J, Hasher L. Is memory schematic? *Psychological Bulletin*. 1983; 93:203–231.
- Anderson N, Barrios A. Primary effects in personality impression formation. *Journal of Abnormal and Social Psychology*. 1961; 63:346–350. [PubMed: 13861274]
- Bayen U, Nakamura G, Dupuis S, Yang C. The use of schematic knowledge about sources in source monitoring. *Memory and Cognition*. 2000; 28(3):480–500. [PubMed: 10881564]
- Boshyan J, Zebrowitz L, Franklin R, McCormick C, Carre J. Age similarities in recognizing threat From faces and diagnostic cues. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*. 2013
- Brownlow S. Seeing is believing: facial appearance, credibility, and attitude-change. *Journal of Nonverbal Behavior*. 1992; 16:101–115.
- Cahill L, McGaugh J. Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences*. 1998; 21(7):294–299. [PubMed: 9683321]
- Carlston DE, Skowronski JJ. Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*. 1994; 66:840–880.
- Carstensen L. Social and emotional patterns in adulthood: support for socioemotional selectivity theory. *Psychology and Aging*. 1992; 7(3):331–338. [PubMed: 1388852]
- Carstensen L, Mikels J. At the intersection of emotion and cognition. *Current Directions in Psychological Science*. 2005; 14(3):117–121.
- Cassidy B, Gutchess A. Social relevance enhances memory for impressions in older adults. *Memory*. 2012; 20(4):332–345. [PubMed: 22364168]
- Cassidy B, Zebrowitz L, Gutchess A. Appearance-based inferences bias source memory. *Memory and Cognition*. 2012; 40(8):1214–1224. [PubMed: 22806429]
- Castle E, Eisenberger N, Seeman T, Moons W, Boggero I, Grinblatt M, Taylor S. Neural and behavioral bases of age differences in perceptions of trust. *Proceedings of the National Academy of Sciences*. 2012; 109(51):20848–20852.
- Chown S. Friendship in old age. *Personal Relationships*. 1981; 2:231–246.
- Erber R, Fiske S. Outcome dependency and attention to inconsistent information. *Journal of Personality and Social Psychology*. 1984; 47(4):709–726.
- Fiske S, Cuddy A, Glick P. Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*. 2007; 11(2):77–83. [PubMed: 17188552]
- Follett K, Hess T. Aging, cognitive complexity, and the fundamental attribution error. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2002; 57B(4):312–323.
- Folstein M, Folstein S, McHugh P. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198. [PubMed: 1202204]
- Franklin R, Zebrowitz L. Older adults' trait impressions of faces are sensitive to subtle resemblance to emotions. *Journal of Nonverbal Behavior*. 2013; 37(3):139–151. [PubMed: 24058225]
- Grady C, Craik F. Changes in memory processing with age. *Current Opinion in Neurobiology*. 2000; 10(2):224–231. [PubMed: 10753795]
- Hansen K, Gerbasi M, Todorov A, Kruse E, Pronin E. People claim objectivity after knowingly using biased strategies. *Personality and Social Psychology Bulletin*. 2014

- Hassin RR, Trope Y. Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*. 2000; 78(5):837–852. [PubMed: 10821193]
- Hedden T, Park D, Nisbett R, Ji L, Jing Q, Jiao S. Cultural variation in verbal versus spatial neuropsychological function across the lifespan. *Neuropsychology*. 2002; 16:65–73. [PubMed: 11853358]
- Hess T. Adaptive aspects of social cognitive functioning in adulthood: Age-related goal and knowledge influences. *Social Cognition*. 2006; 24(3):279–309.
- Hess T, Auman C. Aging and social expertise: The impact of trait-diagnostic information on impressions of others. *Psychology and Aging*. 2001; 16(3):497–510. [PubMed: 11554526]
- Hess T, Follett K. Adult age differences in the use of schematic and episodic information in making social judgments. *Aging and Cognition*. 1994; 1(1):54–66.
- Hess T, Osowski N, Leclerc C. Age and experience influences on the complexity of social inferences. *Psychology and Aging*. 2005; 20(3):447–459. [PubMed: 16248704]
- Hess T, Tate C. Adult age differences in explanations and memory for behavioral information. *Psychology and Aging*. 1991; 6(1):86–92. [PubMed: 2029372]
- Hicks J, Cockman D. The effect of general knowledge on source memory and decision processes. *Journal of Memory and Language*. 2003; 48(3):489–501.
- Kleider H, Cavrak S, Knuycky L. Looking like a criminal: stereotypical black facial features promote face source memory error. *Memory and Cognition*. 2012; 40(8):1200–1213. [PubMed: 22773417]
- Lord C, Ross L, Lepper M. Biased assimilation and attitude polarization: the effect of prior theories on subsequently consider evidence. *Journal of Personality and Social Psychology*. 1979; 37(11):2098–2109.
- Macrae CN, Bodenhausen G, Schloerscheidt A, Milne A. Tales of the unexpected: executive function and person perception. *Journal of Personality and Social Psychology*. 1999; 76(2):200–213. [PubMed: 10074705]
- Marsh R, Cook G, Hicks J. Gender and orientation stereotypes bias source-monitoring attributions. *Memory*. 2006; 14(2):148–160. [PubMed: 16484106]
- Mather M, Carstensen L. Aging and motivated cognition: the positivity effect in attention and memory. *Trends in Cognitive Sciences*. 2005; 9(10):496–502. [PubMed: 16154382]
- Mather M, Johnson M. Affective review and schema reliance in memory in older and young adults. *American Journal of Psychology*. 2003; 116(2):169–189. [PubMed: 12762174]
- Mattarozzi K, Todorov A, Codispoti M. Memory for faces: the effect of facial appearance and the context in which the face is encountered. *Psychological Research*. 2014:1–10.
- May C, Rahhal T, Berry E, Leighton E. Aging, Source Memory, and Emotion. *Psychology and Aging*. 2005; 20(4):571–578. [PubMed: 16420132]
- Mayr S, Erdfelder E, Buchner A, Faul F. A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology*. 2007; 3(2):51–59.
- McCarthy R, Skowronski JJ. What will Phil do next?: Spontaneously inferred traits influence predictions of behavior. *Journal of Experimental Social Psychology*. 2011; 47(2):321–332.
- Mealey L, Daoood C, Krage M. Enhanced memory for faces of cheaters. *Ethology and Sociobiology*. 1996; 17(2):119–128.
- Miner M, Park D. A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*. 2004; 36(4):630–633.
- Mueller J, Heesacker M, Ross M, Nicodemus D. Emotionality of encoding activity in face memory. *Journal of Research in Personality*. 1983; 17(2):198–217.
- Mueller J, Thompson W, Vogel J. Perceived honesty and face memory. *Personality and Social Psychology Bulletin*. 1988; 14(1):114–124.
- Nash R, Bryer O, Schlaghecken F. Look who's talking! Facial appearance can bias source monitoring. *Memory*. 2010; 18(4):451–457. [PubMed: 20408040]
- Nisbett R, Zukier H, Lemley R. The dilution effect: nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*. 1981; 13(2):248–277.
- North M, Fiske S. An inconvenienced youth? Ageism and its potential intergenerational roots. *Psychological Bulletin*. 2012; 138(5):982–997. [PubMed: 22448913]

- Olivola C, Todorov A. Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*. 2010a; 34:83–110.
- Olivola C, Todorov A. Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*. 2010b; 46:315–324.
- Oosterhof N, Todorov A. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*. 2008; 105(32):11087–11092.
- Paivio A. Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*. 1991; 45(3):255–287.
- Peters E, Vastfjall D, Garling T, Slovic P. Affect and decision making: a “hot” topic. *Journal of Behavioral Decision Making*. 2006; 19(2):79–85.
- Plaks J, Stroessner S, Dweck C, Sherman J. Person theories and attention-allocation: Preferences for stereotypic versus counterstereotypic information. *Journal of Personality and Social Psychology*. 2001; 80(6):876–893. [PubMed: 11414372]
- Pratto F, John O. Automatic vigilance: the attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*. 1991; 61(3):380–391. [PubMed: 1941510]
- Radvansky G, Lynchard N, von Hippel W. Aging and stereotype suppression. *Aging, Neuropsychology, and Cognition*. 2009; 16:22–32.
- Rahhal T, May C, Hasher L. Truth and character: sources that older adults can remember. *Psychological Science*. 2002; 13(2):101–105. [PubMed: 11933991]
- Rozin P, Royzman E. Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*. 2001; 5(4):296–320.
- Rudoy J, Paller K. Who can you trust? Behavioral and neural differences between perceptual and memory-based influences. *Frontiers in Human Neuroscience*. 2009
- Ruffman T, Murray J, Halberstadt J, Vater T. Age-related differences in deception. *Psychology and Aging*. 2012; 27(3):543–549. [PubMed: 21463058]
- Rule N, Slepian M, Ambady N. A memory advantage for untrustworthy faces. *Cognition*. 2012; 125:207–218. [PubMed: 22874071]
- Said C, Sebe N, Todorov A. Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*. 2009; 9(2):260–264. [PubMed: 19348537]
- Shi L, Tang W, Liu X. Age-related schema reliance of judgments of learning in predicting source memory. *Aging, Neuropsychology, and Cognition*. 2012; 19:1–2.
- Shipley, W. Shipley Institute of Living Scale. Los Angeles: Western Psychological Services; 1986.
- Skowronski J, Carlston D. Social judgment and social memory: the role of cue diagnosticity in negative, positivity, and extremity biases. *Journal of Personality and Social Psychology*. 1987; 52(4):689–699.
- Skowronski J, Carlston D. Negativity and extremity biases in impression formation: a review of explanations. *Psychological Bulletin*. 1989; 105(1):131–142.
- Somerville L, Wig G, Whalen P, Kelley WM. Dissociable Medial Temporal Lobe Contributions to Social Memory. *Journal of Cognitive Neuroscience*. 2006; 18(8):1253–1265. [PubMed: 16859412]
- Spaniol J, Voss A, Grady C. Aging and emotional memory: Cognitive mechanisms underlying the positivity effect. *Psychology and Aging*. 2008; 23(4):859–872. [PubMed: 19140656]
- Stangor C, McMillan D. Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*. 1992; 111(1): 42–61.
- Stirrat M, Perrett D. Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychological Science*. 2010; 21(3):349–354. [PubMed: 20424067]
- Suzuki A, Suga S. Enhanced memory for the wolf in sheep’s clothing: facial trustworthiness modulates face-trait associative memory. *Cognition*. 2010; 117:224–229. [PubMed: 20804978]
- Todorov A. Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*. 2008; 1124:208–224. [PubMed: 18400932]

- Todorov A, Olson I. Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged. *Social Cognitive and Affective Neuroscience*. 2008; 3:195–203. [PubMed: 19015111]
- Todorov A, Uleman J. Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*. 2002; 83(5):1051–1065. [PubMed: 12416911]
- Todorov A, Uleman J. The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*. 2003; 39(6):549–562.
- Uleman J, Saribay SA, Gonzalez CM. Spontaneous inferences, implicit impressions, and implicit theories. *Annual Review of Psychology*. 2008; 59:329–360.
- Verosky S, Todorov A. Generalization of affective learning about faces to perceptually similar faces. *Psychological Science*. 2010; 21(6):779–785. [PubMed: 20483821]
- Verosky S, Todorov A. When physical similarity matters: Mechanisms underlying affective learning generalization to the evaluation of novel faces. *Journal of Experimental Social Psychology*. 2013; 49:661–669.
- Wagner H. On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*. 1993; 17:3–28.
- Wechsler, D. Wechsler Memory Scale. 3. San Antonio, TX: The Psychological Corporation; 1997.
- Westfall J, Kenny D, Judd CM. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*. in press.
- Willis J, Todorov A. First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*. 2006; 17(7):592–598. [PubMed: 16866745]
- Wood S, Kisley M. The negativity bias is eliminated in older adults: Age related reduction in event-related brain potentials associated with evaluative categorization. *Psychology and Aging*. 2006; 21(4):815–820. [PubMed: 17201501]
- Zebrowitz L, Franklin R, Hillman S, Boc H. Older and younger adults' first impressions from faces: Similar in agreement but different in positivity. *Psychology and Aging*. 2013; 28(1)
- Zebrowitz L, McDonald S. The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*. 1991; 15(6):603–623.
- Zebrowitz L, Montepare J. Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*. 2008; 2/3:1497–1517. [PubMed: 20107613]

Figure 1.

Example congruent, incongruent, and average-face pairs. The terms “congruent,” “incongruent,” and “average-face” were not viewed. (A). Example retrieval stimuli (B).

Figure 2.
Example congruent and incongruent face-behavior pairs performing high and low arousal behaviors.

Figure 3. Congruity effects for trustworthiness- but not dominance-related judgments. Error bars represent standard error. * $p < 0.05$

Figure 4. Congruity effects in young adults' (YA) approach, prosociality, and trustworthiness judgments when pairs involved high arousal behaviors. Congruity effects in older adults' (OA) approach judgments. Error bars represent standard error. * $p < 0.05$

Table 1

Young (YA) and older (OA) adults' unbiased hit rates (M (SD)) for congruent, incongruent, and average-face pairs.

	Congruent	Incongruent	Average-face
YA	0.54 (0.15)	0.52 (0.15)	0.47 (0.17)
OA	0.54 (0.16)	0.50 (0.19)	0.45 (0.19)