Supplemental Materials

Summarized Attribute Preferences Have Unique Antecedents and Consequences

Aline da Silva Frost[1]

Y. Andre Wang[2]

Paul W. Eastwick[1]

Alison Ledgerwood[1]

[1]University of California, Davis

[2]University of Toronto

Table of Contents

Study S1

In this study, we tested whether experimentally manipulating a set of target faces to be more (vs. less) likeable would lead participants to infer stronger summarized preferences for a novel attribute. The procedure was similar to Studies 1 and 2, except that we used both male faces and female faces in this study. Participants evaluated either female faces or male faces based on the sex they were primarily attracted to.

Method

Female faces.

Participants and power. One hundred one participants primarily attracted to women completed the study online through Amazon's Mturk platform. They were randomly assigned to one of two between-subjects conditions (low average likeability vs. high average likeability). We decided a priori to target a cell size of 50 participants based on our lab's standard practice for minimum cell size (the total number of completed surveys in Qualtrics ended up being slightly higher).

We set and recorded the following a priori exclusion criteria: We would exclude participants who (a) gave an identical rating to all faces presented for measurement of functional preferences, and/or (b) provided a nonsensical response to a Winograd-like schema designed to filter out bots or inattentive participants. The numbers of participants who met each of these exclusion criteria were 0 and 3, respectively, resulting in a final sample of N = 98 (11 women, 85 men, and two people who chose another option; M_{age} = 33.7, SD_{age} = 9.7).

Procedure. The procedure was identical to Study 1. First, participants saw a series of 24 faces, each presented along with its level of Reditry, and rated their romantic liking for each pictured person. After the trials, participants completed a measure of their overall summarized preference for Reditry. Lastly, after seeing another survey unrelated to the current research questions, participants completed the attention check and a demographic survey.

New materials and measures. We selected 48 White female faces from the Chicago Faces Database (CFD; Ma et al., 2015). To manipulate average likeability, we divided the faces into two sets of 24 faces that varied similarly in babyfacedness (according to the norming-data ratings in Ma et al., 2015) and that differed only in how likeable they were on average. In a previously published sample (Eastwick & Smith, 2018), 677 participants who were primarily attracted to women evaluated each face on a measure of romantic likeability using 1-7 rating scales. The faces we chose for the high likeability condition had a mean of $M = 3.49$ ($SD = 0.89$) on this scale, and the faces we chose for the low likeability condition had a mean of $M = 2.10$ ($SD = 0.59$). To avoid unintentionally manipulating the strength of the association between babyfacedness and likeability, we ensured that the correlations between the Ma et al. (2015) ratings of babyfacedness and the Eastwick and Smith (2018) ratings of likeability were similar across conditions ($r = .25$ in the high likeability condition and $r = .26$ in the low likeability condition); we also checked that this correlation was similar to the correlation between babyfacedness and likeability in the full population of White female faces in the CFD ($r = .28$). Finally, we inspected the scatterplot between these two variables (see Figure S1).

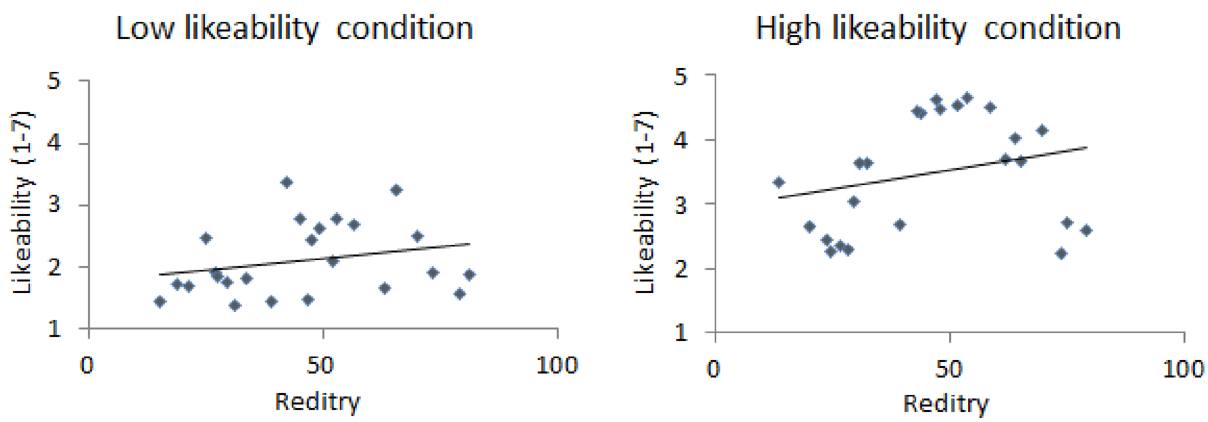

Figure S1. Scatterplots of the female faces used for Study S1. Each dot represents a female face target. Notice that the correlation between pretest ratings of likeability and Reditry was the same in both conditions, whereas the average likeability was higher in the high (vs. low) likeability condition.

Male faces.

Participants and power. One hundred twenty-three participants primarily attracted to males completed the study online through Amazon's Mturk platform. They were randomly assigned to one of two between-subjects conditions (low average likeability vs. high average likeability). The power calculation and exclusion criteria were identical to those described above. The number of participants who gave identical ratings was 3 and the number of participants who provided a nonsensical response to the Winograd-like schema was 3, resulting in a final sample of N = 117 (107 women, 10 men, and 1 person who chose another option; M_{age} = 36.9, SD_{age} = 12.8).

Procedure. The procedure was identical to Studies 1 and 2, as well as the one described above for female faces.

New materials and measures. We used the same male faces as we did in Study 2. As a reminder, the average likeability of faces in the high likeability condition was M = 2.76 (SD = 0.59), and the average likeability of faces in the low likeability condition was

M = 1.83 (*SD* = 0.38). We ensured that the correlation between pretest ratings of babyfacedness and likeability was similar across conditions (*r* = .05 in both conditions; see Figure S2) and reflected the actual correlation in the full population of White male faces in the Chicago Faces Database (*r* = .01).

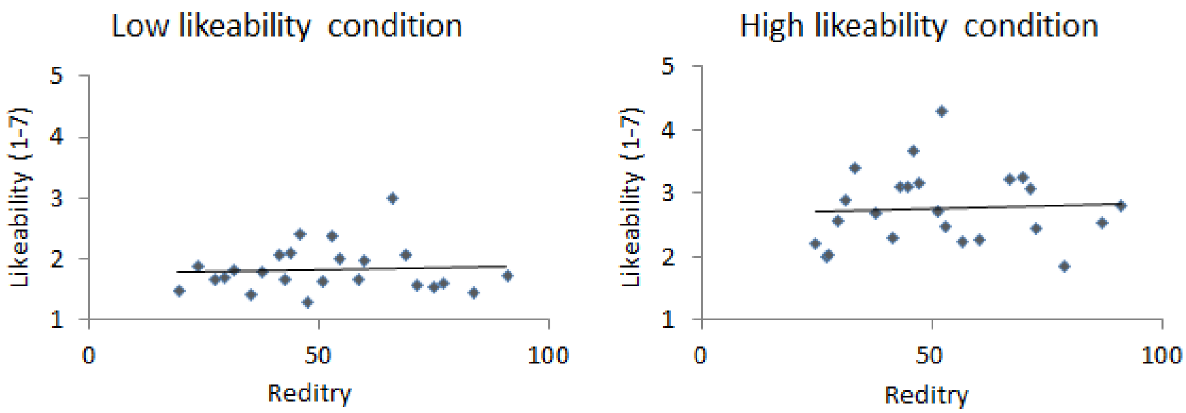

Figure S2. Scatterplots of the male faces used for Study S1. Each dot represents a male face target. Notice that the correlation between pretest ratings of likeability and Reditry was the same in both conditions, whereas the average likeability was higher in the high (vs. low) likeability condition.

Results

Female faces.

Manipulation check. We checked whether the manipulation of average target likeability successfully influenced the amount of liking that participants experienced when learning about their preferences. Our manipulation of average target likeability was successful: On average, participants in the high likeability condition experienced greater liking for the faces they saw (*M* = -0.22, *SD* = 1.47) than participants in the low likeability condition (*M* = -1.65, *SD* = 1.33), *t*(96) = 5.04, *p* < .001, *d* = 1.02, 95% CI [0.60, 1.44].

Functional preferences for Reditry. Although we took care to ensure that the correlation between Reditry and pretest ratings of face likeability was similar across conditions, our manipulation of average likeability unintentionally influenced participants' experienced functional preferences for Reditry, such that their functional preferences for Reditry was higher in the high (vs. low) likeability condition (M = 0.24, SD = 0.30 vs. M = 0.13, SD = 0.16), $t(96)$ = 2.32, p = .022, d = 0.46, 95% CI [0.06, 0.86]. This result underscores the difficulty of selecting appropriate stimuli and the importance of testing different sets of stimuli.

Main analysis. We planned a priori to test the effect of the likeability manipulation on summarized preferences (H1). The results showed that summarized preferences were higher in the high vs. low likeability condition, M = 0.20, SD = 1.68 vs. M = -0.98, SD = 2.27, $t(96)$ = 2.94, p = .004, d = 0.59, 95% CI [0.19, 0.97]. Because our manipulation unintentionally affected functional preferences (see above), our interpretation of this effect is ambiguous: The difference in functional preferences could be driving this effect, instead of the hypothesized difference in likeability. Although this alternative interpretation is unlikely, given that summarized and functional preferences were uncorrelated in this sample, we are hesitant to draw conclusions from this result.

Male faces.

Manipulation check. We checked whether the manipulation of average target likeability successfully influenced the amount of liking that participants experienced when learning about their preferences. Our manipulation of average target likeability was successful: On average, participants in the high likeability condition experienced greater liking for the faces they saw (M = -1.07, SD = 1.35) than participants in the low likeability condition (M = -1.93, SD = 1.22), $t(115)$ = 3.53, p < .001, d = 0.67, 95% CI [0.29, 1.05].

Functional preferences for Reditry. Functional preferences were very similar across the two conditions (M = 0.03, SD = 0.31 vs. M = 0.00, SD = 0.23), t(115) = 0.61, p = .543, d = 0.12, 95% CI [-0.49, 0.25], confirming that our manipulation of average target likeability did not affect participants' functional preferences for Reditry.

Main analysis. After confirming that our manipulation was successful at influencing average liking but not functional preferences for Reditry, we proceeded to our main analysis (H1). Indeed, participants inferred stronger summarized preferences for Reditry in the high versus low likeability conditions (M = -0.58, SD = 1.94 vs. M = -1.34, SD = 1.90), t(115) = 2.11, p = .037, d = 0.40, 95% CI [0.03, 0.77]. In other words, participants inferred that they liked Reditry substantially more when they learned about their preference in a context with high (vs. low) likeability targets.

Study S2

Throughout this manuscript, we use the conceptual term "romantic liking" interchangeably with the terms "romantic interest" and "romantic desire;" these latter two terms were used in the actual items that participants rated in the studies. But are these terms sufficiently highly associated that it is appropriate to treat them synonymously? To test this idea, we asked participants to rate the faces used in Studies 1 and 2 on these three items.

Method

Participants. One hundred nine participants completed the study online through Amazon's Mturk platform. We decided a priori to collect data from at least 100 participants. Consistent with all other studies, we limited participants to those who were between 18 and 35 years old. We set and recorded the following a priori exclusion criteria: We would exclude participants who (a) gave an identical rating to all faces presented, (b) provided a nonsensical response to a Winograd-like schema designed to filter out bots or inattentive participants, or (c) expressed suspicion about the purpose of the study. One participant met the second exclusion criterion and was excluded, resulting in a final sample of N = 108 (56 women, 51 men, and 1 person who chose another option; M_{age} = 28.1, SD_{age} = 4.0).

Procedure. Participants first completed a brief prescreen in which they indicated their age and the sex to which they were primarily romantically attracted, which determined the sex of the potential partners presented to them throughout the rest of the study. Participants then saw a series of 48 faces; the female faces were the same as those presented in Study 1, and the male faces were the same as those presented in Study 2. Participants rated each pictured person on romantic interest ("To what extent are you romantically interested in this person?" -4 = *strongly dislike*, -4 = *strongly like*), romantic liking ("To what extent do you romantically like this person?" -4 = *strongly dislike*, 4 = *strongly like*), and romantic desire ("To what extent do you experience romantic desire for this person?" 1 = *not at all*, 9 = *a great deal*), all on 9-point Likert-type scales. The question on romantic interest was worded to exactly match that of the measure of functional preferences in Studies 1–3, and the question on romantic desire was worded to exactly match that of the measure of functional preferences in Study 4. Lastly, participants completed the same attention check as used in Studies 1–3, a demographic survey, and a short questionnaire unrelated to the current research questions.

Results and Discussion

As primary analyses, we conducted cross-classified mixed effects modeling to test the relation between ratings on each pair of variables (romantic interest, romantic liking, and romantic desire). Because each participant rated each variable across faces, cross-classified mixed effects models allowed us to distinguish between the fixed effects among the variables (e.g., the relation between interest and liking) and the random effects of stimuli and participants (Raudenbush & Bryk, 2002). Accounting for the nested nature of the data at both the participant and the stimulus level also provided us with more accurate estimates (Judd et al., 2012, 2017). We ran three models: one in which desire was regressed on interest, one in which interest was regressed on liking, and one in which liking was regressed on desire. In addition to the fixed effect of the independent variable (IV) on the dependent variable (DV), each model also included random intercepts by participant and by stimuli. Ratings on each variable were standardized across faces and participants (i.e., grand-mean centered). The models were expressed in R as:

DV ~ IV + (1 | participant) + (1 | stimulus)

Results showed that the three variables were highly associated with each other. Romantic interest significantly predicted romantic desire, $\beta = .87$, 95% CI [.86, .89], romantic liking significantly predicted romantic interest, $\beta = .93$, 95% CI [.92, .94], and romantic desire significantly predicted romantic liking, $\beta = .84$, 95% CI [.82, .86], all *p*s < .001.

We also conducted secondary analyses in which we calculated between-subjects correlations among the three variables by averaging ratings by each participant across faces, and the results were highly similar: $r_{\text{desire.interest}} = .88$, 95% CI [.82, .92], $r_{\text{interest.liking}} = .99$, 95% CI [.98, .99], $r_{\text{liking.desire}} = .86$, 95% CI [.80, .91], all *p*s < .001. Therefore, we concluded that participants' ratings on romantic interest, romantic liking, and romantic desire were interchangeable.

Study S3

In Study 4, we used photographs that we collected for use as stimuli from a publicly accessible dating website. Following the stimuli collection procedure of Wood and Brumbargh (2009), we set three criteria for the selected photographs, which were required to (a) show at least the person's head and torso in full view, (b) be of a reasonably high quality (i.e., not blurry or unfocused, or so small that facial features cannot be discerned), and (c) contain only one individual. Consistent with Wood and Brumbargh (2009), we selected all photographs from the "aged 18 to 25" range on the website and stopped stimuli collection once 100 photographs per target sex met our criteria to ensure a random cross-section of photographs.

Following stimuli collection, we conducted a separate rating study. Participants (*N* = 132; 71 women, 61 men) between 18 and 35 years old (*M* = 28.8, *SD* = 4.0) completed the study online through Mturk. Participants rated each of the 100 preferred-sex targets on a list of attributes that people can rapidly and consensually rate on from faces (Oosterhof & Todorov, 2008). On each screen, participants saw one target and the list of attributes below the target's photograph, and participants rated the target on each attribute on a 9-point scale (1 = *not at all*, 9 = *extremely*). Cronbach's alphas were high for ratings on both "intelligent" (α = .95 for male targets and α = .86 for female targets) and "confident" (α = .94 for male targets and α = .88 for female targets; see Table S1 for the full list of attributes rated and their descriptive statistics).

Table S1

*Internal Consistency (Cronbach's Alpha) and Interrater Agreement (*r*) of Ratings on Attributes*

Attribute	Cronbach's alpha (α)		Interrater agreement (*r*)	
	Male Targets	Female Targets	Male Targets	Female Targets
Attractive	.96	.97	.27	.28
Mean	.93	.80	.17	.07
Dominant	.94	.84	.18	.08
Trustworthy	.94	.80	.17	.06
Aggressive	.94	.81	.18	.06
Caring	.93	.78	.18	.06
Emotionally stable	.93	.84	.18	.07
Responsible	.95	.85	.21	.08
Sociable	.93	.86	.18	.09

Confident	.94	.88	.21	.10
Intelligent	.95	.86	.21	.08
Sensitive	.93	.73	.16	.04

Note: N = 66 for ratings of 100 male targets, and N = 66 for ratings of 100 female targets.

Study 4: Additional Results

Hypotheses 3 and 4: Preregistered Analyses

Primary situation-selection dependent measures.

Table S2.

Fit Indices from Structural Equation Models with Summarized and Functional Preferences Predicting Primary Dependent Variables in Study 4.

Predictor Type	Dependent Variables	Attributes	χ^2	*df*	*p*	CFI	TLI	RMSEA
SP	SS at a distance	Intelligence	10.34	7	.170	1.00	1.00	0.03
		Confidence	8.96	7	.256	1.00	1.00	0.02
FP	SS at a distance	Intelligence	265.30	21	< .001	0.90	0.84	0.14
		Confidence	263.68	21	< .001	0.90	0.84	0.14
SP	SS with experience	Intelligence & Confidence	1.23	7	.990	1.00	1.01	0.00
FP	SS with experience	Intelligence & Confidence	125.86	21	< .001	0.96	0.94	0.09

Note: SP = summarized preferences, FP = functional preferences, SS = situation selection. In all analyses, preferences for both intelligence and confidence were entered as predictors. Because situation selection with experience was a dichotomous choice between two situations (i.e., website with highly intelligent targets and website with highly confident targets), the summarized preference model and the functional preference model estimated the effects of preferences for intelligence and confidence on this DV simultaneously. Therefore, we report only one set of fit indices for each predictor type on that variable. For analyses involving the dichotomous dependent variable (i.e., situation selection with experience), fit indices were calculated using the diagonally weighted least squares (DWLS) estimator.

Secondary situation-selection dependent measures. To examine whether the observed double dissociation between summarized and functional preferences could have been driven by an incidental feature of the format of our primary dependent measures, we conducted planned analyses on our secondary dependent measures using the same analytic approaches. First, we asked whether summarized preferences would still strongly predict situation selection at a distance if we forced a tradeoff between one website versus another. After all, the double dissociation observed in our primary analyses could be driven by a difference in how people responded to a single situation (as measured by the primary situation selection at a distance variable above) rather than a tradeoff between two situations (as measured by the primary situation selection with experience variable above). We examined this possibility by assessing how strongly summarized and functional preferences predicted participants' interest in one website *versus* the other on a bipolar rating scale (i.e., the website that would provide access to highly intelligent partners versus the website that would provide access to highly confident partners; see Table S3 for a summary of effect sizes, and Table S4 for the relevant fit indices; all models fit the data well).

Table S3.

Effect Sizes for Summarized and Functional Preferences Predicting Secondary Dependent Variables in Study 4.

Analytic Approaches	Predictor Type	Dependent Variables	Attributes	
			Intelligence	Confidence
Structural equation models	SP	SS at a distance (tradeoff)	.48***	.38***
	FP	SS at a distance (tradeoff)	.20***	.13**
	SP	SS at a distance (choice)	.28***	.26***
	FP	SS at a distance (choice)	.12**	.12**
Bivariate regression	SP	SS at a distance (tradeoff)	.31***	.11*
	FP	SS at a distance (tradeoff)	.15***	-.04
	SP	SS at a distance (choice)	.12***	.09***
	FP	SS at a distance (choice)	.18	.13
Multiple regression	SP	SS at a distance (tradeoff)	.41***	.29***
	FP	SS at a distance (tradeoff)	.16***	.08
	SP	SS at a distance (choice)	.25***	.23***
	FP	SS at a distance (choice)	.45**	.39*

Note: SP = summarized preferences, FP = functional preferences, SS = situation selection. *p < .05, **p < .01, ***p < .001. All effect sizes are reported in correlation coefficients.

Results from our focal SEM approach revealed that summarized preferences strongly predicted this tradeoff version of the measure (intelligence: b = 1.27, SE = 0.12, p < .001, r = .48, 95% CI [.40, .57]; confidence: b = 0.98, SE = 0.12, p < .001, r = .38, 95% CI [.29, .46]). That is, when we equated our situation selection at a distance and situation selection with experience measures in terms of both forcing a tradeoff, summarized preferences still strongly predicted situation selection at a distance (Table S3), whereas summarized preferences weakly predicted situation selection with experience (see Table 3 in the manuscript). The results were similar across the two traits, and mostly similar across the two alternative analytic approaches.

Table S4.

Fit Indices from Structural Equation Models with Summarized and Functional Preferences Predicting Secondary Dependent Variables in Study 4.

Predictor Type	Dependent Variables	Attributes	χ^2	*df*	*p*	CFI	TFI	RMSEA
SP	SS at a distance (tradeoff)	Intelligence & Confidence	7.81	7	.350	1.00	1.00	.01
	SS at a distance (choice)	Intelligence & Confidence	2.55	7	.923	1.00	1.01	.00
FP	SS at a distance (tradeoff)	Intelligence & Confidence	264.36	21	< .001	0.91	0.84	.14
	SS at a distance (choice)	Intelligence & Confidence	53.89	21	< .001	0.99	0.98	.05

Note: SP = summarized preferences, FP = functional preferences, SS = situation selection, CFI = comparative fit index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation. Because situation selection with experience was a dichotomous choice between two situations (i.e., website with highly intelligent targets vs. website with highly confident targets), one model simultaneously estimated the effects of preferences for both attributes for each predictor type. Therefore, we report only one set of fit indices for each predictor type on that variable. For all analyses, fit indices were calculated using the DWLS estimator.

We also explored whether forcing a tradeoff affected the predictive power of functional preferences. The effects of functional preferences on the tradeoff measure were more ambiguous (Table S3). Across the different attributes and analytic approaches, some of the effect sizes on the tradeoff measure were more similar to those of functional preferences on our original measure of situation selection at a distance, and some of the effect sizes on the tradeoff measure were more similar to those of functional preferences on situation selection with experience (see Table 3 in the manuscript). These intermediate results suggest that there may be something special about tradeoffs that gives functional preferences a little extra predictive power. However, because the results from functional preferences predicting the tradeoff measure were less consistent across the two attributes and alternative analytic approaches, it would be important to replicate these results and find a more consistent pattern across analyses and attributes before drawing any strong conclusions from them.

Next, we asked whether summarized preferences would still strongly predict participants' situation selection at a distance if we asked participants to make a binary choice between one website versus another. After all, the double dissociation observed in our primary analyses could be driven by a difference between using a rating scale to evaluate situation desirability (as measured by our primary situation selection at a distance variable) and making a binary choice between situations (as measured by our primary situation selection with experience variable). We examined this possibility by assessing how strongly summarized and functional preferences predicted participants' choice between the two described websites. Results from our focal SEM approach revealed that summarized preferences still strongly predicted this choice version of the "at a distance" measure (intelligence: $b = 1.04$, $SE = 0.10$, $p < .001$, $r = .28$, 95% CI [.23, .32]; confidence: $b = 1.00$, $SE = 0.13$, $p < .001$, $r = .26$, 95% CI [.20, .33]). That is, when we equated our situation selection at a distance and situation selection with experience measures in terms of both involving a binary choice, summarized preferences still strongly predicted situation selection at a distance (Table S3), whereas recall that summarized preferences weakly predicted situation selection with experience (see Table 3 in the manuscript). The results were similar across the two attributes and across the alternative analytic approaches.

We can also explore whether using a binary choice version of this "at a distance" measure affected the predictive power of functional preferences. The results from our focal SEM approach revealed that functional preferences weakly predicted the binary version of situation selection at a distance (intelligence: $b = 0.44$, $SE = 0.16$, $p = .005$, $r = .12$, 95% CI [.04, .20]; confidence: $b = 0.45$, $SE = 0.16$, $p = .004$, $r = .12$, 95% CI [.04, .21]). In other words, when we equate our situation selection at a distance and situation selection with experience measures in terms of both involving a binary choice, functional preferences still weakly predicted situation selection at a distance (Table S3), whereas recall that functional preferences strongly predicted situation selection with experience (see Table 3 in the manuscript). The results were similar across the two attributes and mostly similar across the alternative analytic approaches.[1]

Exploratory analyses on retaining the item "charismatic" in summarized preferences for confidence. We explored the impact of retaining the item "charismatic" in our calculation of summarized preference for confidence on predictions by summarized preferences. The results did not change in any substantive way (see Table S5). In addition, all five structural equation models with summarized preferences as predictors fit the data at a level comparable with their corresponding models reported in the manuscript: summarized preference for intelligence predicting situation selection at a distance, $\chi^2(12) = 21.20$, $p = .047$, CFI = 0.99, TLI = 0.99, RMSEA = .04; summarized preference for confidence predicting situation selection at a distance, $\chi^2(12) = 24.41$, $p = .018$, CFI = 0.99, TLI = 0.99, RMSEA = .04; summarized preferences predicting situation selection with experience, $\chi^2(12) = 6.11$, $p = .911$, CFI = 1.00, TLI = 1.01, RMSEA = .00; summarized preferences predicting situation selection with experience (tradeoff), $\chi^2(12) = 18.74$, $p = .095$, CFI = 1.00, TLI = 0.99, RMSEA = .03; summarized preferences predicting situation selection (choice), $\chi^2(12) = 6.76$, $p = .873$, CFI = 1.00, TLI = 1.00, RMSEA = .00.

[1] Note that it is complicated to compare the strength of the predictive power of summarized versus functional preferences directly. For example, although the effect sizes for functional preferences predicting situation selection at a distance (choice) from the bivariate regressions and multiple regressions were nominally larger than those for summarized preferences, they were less significant (i.e., their associated *p* values were larger) due to greater uncertainty around their estimates. In contrast, one can more readily compare the magnitude of the coefficients for summarized preferences predicting different dependent measures, and the magnitude of the coefficients for functional preferences predicting different dependent measures, due to comparable standard errors.

Table S5.

Test Statistics and Effect Sizes for Summarized Preferences Predicting Primary and Secondary Dependent Variables, with Three Indicators of Summarized Preference for Confidence.

Analytic Approaches	Dependent Variables	Attributes									
		Intelligence					Confidence				
		b (*SE*)	*p*	*β*	*OR*	*r*	*b* (*SE*)	*p*	*β*	*OR*	*r*
Structural equation models	SS at a distance	0.89 (0.13)	< .001	0.39	-	.32	0.89 (0.13)	< .001	0.39	-	.32
	SS with experience	0.45 (0.13)	< .001	-	1.56	.12	0.52 (0.13)	< .001	-	1.69	.14
	SS at a distance (tradeoff)	1.30 (0.12)	< .001	0.60	-	.49	1.02 (0.12)	< .001	0.47	-	.39
	SS at a distance (choice)	1.10 (0.10)	< .001	-	3.02	.29	1.05 (0.13)	< .001	-	2.85	.28
Bivariate regression	SS at a distance	0.73 (0.09)	< .001	0.33	-	.33	0.68 (0.09)	< .001	0.32	-	.32
	SS with experience	0.13 (0.10)	.192	-	1.13	.03	0.20 (0.09)	.020	-	1.23	.06
	SS at a distance (tradeoff)	0.66 (0.09)	< .001	0.31	-	.31	0.24 (0.09)	.005	0.12	-	.12
	SS at a distance (choice)	0.44 (0.09)	< .001	-	1.55	.12	0.33 (0.10)	.001	-	1.39	.09
Multiple regression	SS at a distance	0.73 (0.10)	< .001	0.33	-	.29	0.75 (0.10)	< .001	0.35	-	.30
	SS with experience	0.31 (0.11)	.005	-	1.37	.09	0.36 (0.10)	< .001	-	1.43	.10
	SS at a distance (tradeoff)	1.04 (0.09)	< .001	0.49	-	.43	0.75 (0.09)	< .001	0.36	-	.32
	SS at a distance (choice)	0.98 (0.14)	< .001	-	2.67	.26	0.93 (0.14)	< .001	-	2.53	.25

Note: SS = situation selection. Unstandardized regression coefficients (*b*) for dichotomous variables are logit coefficients.

Correlation between Functional and Summarized Preferences

Study S1

The correlation between functional and summarized preferences for Reditry was $r = .06$, $p = .532$, 95% CI [-.14, .26] for female faces and $r = .17$, $p = .067$, 95% CI [-.01, .34] for male faces.

Study 1

The correlation between functional and summarized preferences for Reditry was $r = .11$, $p = .309$, 95% CI [-.10, .30].

Study 2

The correlation between functional and summarized preferences for Reditry was $r = .09$, $p = .259$, 95% CI = [-.06, .24].

Study 3

The correlation between functional and summarized preferences for Reditry was $r = .11$, $p = .008$, 95% CI [.03, .19].

Meta-analysis

Meta-analytical results of the correlations between functional and summarized preferences for Reditry in Studies 1-3 and Study S1 were $N = 1{,}046$, $r = .12$, $z = 3.74$, $p < .001$, 95% CI [.06, .18].

Study 4

Note that the large sample size employed in this study exceeds Schönbrodt and Perugini's (2014) minimum recommendation for stable effect size estimates, and so we can have a relatively high degree of confidence in the stability of the observed correlations. The correlation between functional and summarized preferences for intelligence was $r = .18$, $p < .001$, 95% CI

[.10, .26], and the correlation between functional and summarized preferences for confidence was $r = .08$, $p = .045$, 95% CI [.002, .17]. Notably, the CIs for both attributes were compatible with the CIs observed for Reditry in our previous studies.

References

Eastwick, P. W., & Smith, L. K. (2018). Sex-differentiated effects of physical attractiveness on romantic desire: A highly powered, preregistered study in a photograph evaluation context. *Comprehensive Results in Social Psychology*, *3*, 1-27.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54-69.

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*, 601-625.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*, 1122-1135.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, USA, 105*, 11087-11092.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Atlanta, GA: Sage.

Wood, D., & Brumbaugh, C. C. (2009). Using revealed mate preferences to evaluate market force and differential preference explanations for mate selection. *Journal of Personality and Social Psychology*, *96*, 1226-1244.