**Table of Contents**

Parallels between PERSON and MET.…………………………..………………………… p. 2

Cumulative Feature Lens Variance in Two Speed-dating Samples...……………………… p. 3

Calculating the Effects of the Four Lenses as “Weights” with a Full SRM Design……….. p. 6

Cumulative Effects of Sex Differences in Body Features in a Speed-dating Study.……… p. 9

**Parallels between PERSON and MET**

The four sources of information in MET draw inspiration from the six components of the PERSON model (Kenny, 2004). As discussed previously, the PERSON model operates upstream of MET; it describes how people translate acts in the behavior stream (e.g., “Kevin is smiling”) into semantic judgments (e.g., “Kevin is friendly”) by drawing from different sources of information in the mind (e.g., stereotypes, personality, norms). The PERSON model’s six sources of information map onto the four sources we use here, albeit indirectly. The differences are due to the fact that (a) PERSON is designed to distinguish between appearance and behavioral cues (a distinction that is not of primary interest to MET), and (b) PERSON does not emphasize the distinction between perceiver and relationship variance (which is critical to MET). Nevertheless, the general principles are similar.

For the interested reader, the correspondence between the two models is as follows: The common lens corresponds to the PERSON model components S (stereotype), P (personality), and the portion of N (norm) that applies to acts that all perceivers have witnessed (which will be very small in real romantic relationships); the perceiver lens corresponds to the mean of R (residual) and O (opinion); and the feature and target-specific lenses correspond to variance in R, O, E (error), and the (large) portion of N that applies to acts that only the perceiver has witnessed. Another important difference is that, in MET, all components reside in the perceiver’s mind, but P in PERSON is actually a force that resides within the target him/herself.

**Cumulative Feature Lens Variance in Two Speed-dating Samples**

In the section “The Four Sources in the Information Store” in the main manuscript, we described that a third (rarely examined) conceptualization of the feature lens is the extent to which there is random variability in the association of attributes with an evaluative outcome. That is, are there consistent individual differences in the association of an attribute (e.g., physical attractiveness) with romantic evaluations in the first place? Such random variability would typically be considered a precursor to the common attribute-matching approaches to testing the feature lens. In other words, there must be some random variability in the association of physical attractiveness with a romantic evaluation in order for an individual difference (e.g., people’s ideal partner preferences for attractiveness) to exhibit a significant attribute-matching effect (e.g., an ideal partner preference × attribute interaction).

To try to estimate the random variability in the extent to which a set of attributes predicts a romantic evaluation, we drew from two of our own speed-dating datasets (NSDS-I *N* = 163; Eastwick & Finkel, 2008; NSDS-II *N* = 187, Tidwell et al., 2013, Northwestern University IRB 1343-019). Specifically, we ran the following multi-level regression equations using the items in Table S1:

Level 1:

RomanticEvaluation = β0 + β1PhysicalAttractiveness+ β2EarningPotential+ β3Responsive+ β4Vitality+ β5Intelligent+ ε

Level 2:

β0 = γ0 + μ0

β1 = γ1 + μ1

β2= γ2+ μ2

β3 = γ3 + μ3

β4= γ4 + μ4

β5 = γ5 + μ5

These five attributes cover the most central traits that people describe in an ideal romantic partner, drawing from Fletcher et al. (1999). Our approach expands on Fletcher et al. (1999) in that we (a) separate physical attractiveness and vitality into two separate groupings, and (b) include intelligence as its own unique attribute.

The question we tested was: To what extent do the variance and covariance components associated with the five attributes (i.e., σ2μ1, σ2μ2, σ2μ3, σ2μ4, and σ2μ5) collectively account for variance in the dependent measure, as a fraction of the total variance? To calculate this value, we used the approach recommended by Rights and Sterba (2019) and the R package “r2mlm” (Shaw et al., 2020), which provides the percentage of variance accounted for by all random effects components (i.e., “slope variation” or R2t (v)) as a fraction of the total variance. All five independent variables were person centered as recommended by Rights and Sterba (2019). To remove interdependence caused by the fact that men and women evaluated each other at the speed-dating events, we conducted the analyses for men and women separately.

According to this analysis, these attributes collectively accounted for 3.0% (men) and 3.4% (women) of the variance in NSDS-I and 3.5% (men) and 2.8% (women) of the variance in NSDS-II. In other words, all individual differences in the associations of these attributes with a romantic evaluation accounted for ~3% of the variance across all of participants’ speed-dates. These findings imply that the sum of all similarity, ideal partner preference-matching, and other plausible feature lens effects involving these five sets of attributes would be (collectively) quite small.

These estimates are consistent with Prediction #1 in the main manuscript: Feature lens effects are collectively larger than zero (we posit 10% as a “ceiling” in the manuscript), but they are likely small, and they will not be able to explain the bulk of the variance associated with compatibility (i.e., relationship variance).

Sample code for r2mlm:

model\_lme4 <- lmer(RomanticEvaluation ~ 1 + PhysicalAttractiveness + EarningPotential + Responsive + Vitality + Intelligent + (1 + PhysicalAttractiveness + EarningPotential+ Responsive + Vitality + Intelligent | user\_id), data = nsds1\_men, REML = TRUE, control = lmerControl(optimizer = "bobyqa"))

r2mlm(model\_lme4)

Table S1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Construct |  | NSDS-I items | α |  | NSDS-II items | α |
| Independent variable: |  |  |  |  |  |  |
|  Physical attractiveness |  | physically attractive, sexy / hot | .95 |  | physically attractive, sexy / hot | .95 |
|  Earning potential |  | good career prospects,ambitious / driven | .86 |  | good career prospects,ambitious / driven | .88 |
|  Responsive |  | responsive, dependable / trustworthy, friendly / nice | .81 |  | responsive, dependable / trustworthy, friendly / nice | .83 |
|  Vitality |  | fun / exciting, funny, charismatic, confident | .91 |  | fun / exciting, funny, charismatic, confident, assertive | .91 |
|  Intelligent |  | intelligent |  |  | smart, intellectually sharp | .92 |
| Dependent variable: |  |  |  |  |  |  |
|  Romantic evaluation |  | I really liked my interaction partner; I was sexually attracted to my interaction partner; I am likely to say “yes” to my interaction partner | .88 |  | I really liked my interaction partner; I was sexually attracted to my interaction partner; I am likely to say “yes” to my interaction partner | .87 |

Note: Participants completed all items on 1-9 response scales after each speed-date.

**Calculating the Effects of the Four Lenses as “Weights” with a Full SRM Design**

It is perhaps clarifying for some readers to conceptualize the effects of the four lenses as regression weights: Information in memory can be categorized into the four lenses, and when the information in the four lenses is brought to mind, they operate as weights that are applied to the IV 🡪 DV association. In other words, the extent to which a given semantic concept produces a positive evaluation is determined by the application of the CL, PL, FL, and TspL weights. Figure S1 presents a design that illustrates this idea, inspired by an approach used by Lutz and Lakey (2001).

**Figure S1 – Algebraically Separating Weights Deriving from All Four Lenses**



Imagine a researcher were interested in calculating the weights for a semantic concept like *expressing empathy* in a blocked design, where all men evaluate all women (and vice versa). Also imagine that the researcher had collected perceivers’ perceptions of these multiple targets, over multiple moments in time, which would yield assessments of the extent to which the presence vs. absence of the attribute *expressing empathy* affected perceivers’ evaluations of each target. In this case, the researcher could mathematically separate weights for all four lenses for the attribute. Specifically, they could calculate (a) each perceiver’s mean evaluation across all targets and time points (a0: the perceiver lens weight), (b) all perceivers’ average association between *expressing empathy* and their evaluation of the targets (a1: the common lens weight), (c) a given perceiver’s personal association between *expressing empathy* and the evaluation, above and beyond the common lens weight (a2: the feature lens weight), and (d) a given perceiver’s personal association between *expressing empathy* and the evaluation for each target, above and beyond the common and feature lens weights (a3: the target-specific lens weight).

With the weights (from a real dataset) in hand, researchers could address a wide variety of research questions—especially if the weights can be calculated across several attributes (not just *expressing empathy*, but also *asking questions* and *telling funny stories*). For example, different attributes might show more vs. less variability depending on the lens: The feature lens weights might vary more for *expressing empathy* than *telling funny stories* (i.e., perceivers differ substantially in the extent to which they generally like targets who express empathy, but they do not differ substantially in the extent to which they generally like targets who tell funny stories). Furthermore, the target-specific lens weights could show the reverse pattern (i.e., perceivers like some targets, but not others, when they are telling funny stories). In other words, this design could be used to test hypotheses about which attributes tend to be more or less desirable in general (i.e., the common lens), desirable for some perceivers but not others (i.e., the feature lens), or desirable in the context of certain relationships but not others (i.e., the target-specific lens)—all while separating out participants’ general liking tendencies (i.e., the perceiver lens). (See also Lutz and Lakey, 2001, for additional inspiration.) Approaches like this one would allow researchers to make nomothetic inferences about a sample of participants while also using a design that is sufficiently idiographic that it can capture the target-specific lens, which (by definition) varies from relationship to relationship.

**Cumulative Effects of Sex Differences in Body Features in a Speed-dating Study**

In the section “Prediction #3,” we discuss how sex × attribute interaction effects will tend to be small, even for attributes that should theoretically have different evaluative consequences for men and women. We also noted how this prediction applies primarily when the DV is an overall evaluative judgment (rather than something closer to a semantic concept). Support for these two predictions can be found in a recent study by Sidari et al. (2021), who graciously provided their data for the additional analysis described here.

Sidari et al. (2021) collected data on several (“objectively” measured) features of men’s and women’s bodies at a speed-dating event: Shoulders, waist, hips, height, shoulders × waist (traditionally considered more important in men’s bodies), and waist × hips (considered more important in women’s bodies; see Sidari et al., 2021, Supplemental Table 5). To calculate the sum total of the variance accounted for by sex differences in these features, we can again use the approach recommended by Rights and Sterba (2019) and the R package “r2mlm” (Shaw et al., 2020). In this case, we calculate the fixed effects associated with sex differences in all six of these body features by comparing the R2t (f2) (i.e., variance for level-2 fixed effects) values for two models. The first model calculates the extent to which a person’s overall likeability as rated by each opposite-sex speed-dating attendee (i.e., “RomanticEvaluation”) is a function of his/her sex and his/her body features.

Model 1 (no sex differences):

model\_lme4 <- lmer(RomanticEvaluation ~ 1 + sex + height + shoulders + waist + hips + shoulders\*waist + waist\*hips + (1 | ID), data = SidariSD, REML = TRUE, control = lmerControl(optimizer = "bobyqa"))

r2mlm(model\_lme4)

The R2t (f2) value for this analysis is .042; that is, 4.2% of the variance in the romantic evaluation is accounted for by these 7 fixed effects (i.e., sex and six body features). Next, we calculate a model that adds sex differences in all six body features:

Model 2 (all sex differences):

model\_lme4 <- lmer(RomanticEvaluation ~ 1 + sex + height + shoulders + waist + hips + shoulders\*waist + waist\*hips + sex\*height + sex\*shoulders + sex\*waist + sex\*hips + sex\*shoulders\*waist + sex\*waist\*hips + (1 | ID), data = SidariSD, REML = TRUE, control = lmerControl(optimizer = "bobyqa"))

r2mlm(model\_lme4)

The R2t (f2) value for this analysis is .068; that is, 6.8% of the variance in the romantic evaluation is accounted for by these fixed effects. **The difference between the two models is 2.6%**, which implies that the sum of all sex differences in these bodily features explains 2.6% of the variance in romantic evaluations—a small value that is consistent with Prediction #3.

Sidari et al. (2021) also included a measure of “body attractiveness,” which is closer to a semantic concept in MET (although still highly evaluative, of course). If body attractiveness is used instead of overall attractiveness, the value for model 1 is .095 (9.5%) and the value for model 2 is .146 (14.6%), and **the difference between the two models is 5.2%.**

We can conclude two things from these analyses:

1. The sum total of the effects of sex differences of body features on romantic evaluations is (perhaps surprisingly) small (2.6%).
2. Sex differences in the effects of body features are approximately twice as strong for body attractiveness judgments (5.2%) than for overall evaluative judgments (2.6%), which is consistent with the idea in MET that external cues operate upstream of semantic concepts, which operate upstream of evaluative consequences.