

Independent contributions of the face, body and gait to the representation of the whole person

Noa Simhi¹ and Galit Yovel^{1,2}

¹ The School of Psychological Sciences and ² The Sagol School of Neuroscience, Tel-Aviv University, Tel-Aviv 69978, Israel

Correspondence regarding this manuscript should be addressed to:

Galit Yovel and Noa Simhi

School of Psychological Sciences & Sagol School of Neuroscience

Tel Aviv University

Tel Aviv, 69978, Israel

Email: gality@post.tau.ac.il, simhinoa@gmail.com

Tel: 972-3-640-5474

Fax: 972-3-640-9547

Abstract

Most studies on person perception have primarily investigated static images of faces. However, real-life person perception involves also the body and often the gait of the whole person. Whereas some studies indicated that the face dominates the representation of the whole person, others have emphasized the additional contribution of the body and gait. Here, we compared models of whole person perception by asking whether a model that includes the body for static whole person stimuli and also the gait for dynamic whole person stimuli accounts better for the representation of the whole person than a model that takes into account the face alone. Participants rated the distinctiveness of static or dynamic displays of different people based on either the whole person, face, body, or gait. By fitting a linear regression model to the representation of the whole person based on the face, body and gait, we revealed that the face and body contribute uniquely and independently to the representation of the static whole person, and that gait further contributes to the representation of the dynamic person. A complementary analysis examined if these components are also valid dimensions of a whole person representational space. This analysis further confirmed that the body in addition to the face, as well as the gait are valid dimensions of the static and dynamic whole person representations, respectively. These data clearly show that whole person perception goes beyond the face and is significantly influenced by the body and gait.

Introduction

Most studies on person perception presented participants with static images of faces alone. Nonetheless, in naturalistic encounters with people, we see not only their face but also their body, and often we perceive their motion pattern as well. Do the body and motion contribute to person recognition beyond the face? Earlier studies have proposed that the face dominates person perception with little contribution to the body and gait (Burton, Wilson, Cowan, & Bruce, 1999), however in recent years, it has been shown that, under certain circumstances, the body and motion may also contribute to person recognition (see for example Hahn, O’Toole, & Phillips, 2015; O’Toole et al., 2011; Pilz & Thornton, 2017; Pilz, Vuong, Bülthoff, & Thornton, 2011; Simhi & Yovel, 2016, 2017, 2019 for review see Yovel & O’Toole, 2016 and Hu, Baragchizadeh, & O’Toole, 2020).

Studies using static images have demonstrated that the body contributes to person recognition even without people being aware of it, and when subjective reports indicate that participants believe they are basing their judgments on the face (Rice, Phillips, Natu, An, & O’Toole, 2013). The body has been shown to contribute not only to person recognition, but to other types of perceptual decisions involving the whole person – such as judgment of emotion and personality traits. For example, Aviezer, Trope, & Todorov, 2012a demonstrated that body context plays an important role in differentiating between positive and negative extreme emotions, and other studies have shown that faces and bodies are processed holistically when judging emotions (for example Aviezer, Trope, & Todorov, 2012b; Meeren, van Heijnsbergen, & de Gelder, 2005; Van den Stock, Righart, & de Gelder, 2007). Using advanced modeling tools, recently body shape features such as weight or having a rectangular/pear shaped body, have

been shown to predict body-based personality trait inferences as well (Hu, Parde, Hill, Mahmood, & O'Toole, 2018).

The aforementioned studies presented static images of the whole person. Other studies that presented the whole person in motion, have shown that the body and body motion also contribute significantly beyond the face to person recognition (O'Toole et al., 2011; Pilz & Thornton, 2017; Simhi & Yovel, 2016, 2017, 2019). Interestingly, presenting participants with movies of approaching bodies improves face-based person recognition, even when the bodies themselves are not informative – a contribution which is not found for receding or static bodies (Pilz et al., 2011). O'Toole et al., 2011, used an innovative fusion approach in order to estimate computationally the optimal person recognition strategy in a task which included static and dynamic stimuli depicting the whole person, face alone or body alone. They found that in recognition from static images the optimal recognition strategy encourages reliance on the face, while recognition from dynamic displays is best when taking into account information from both the face and body. Consistent with these findings, Simhi & Yovel, 2016 have shown that the body contributes to whole person perception beyond the face for dynamic but not static whole person stimuli. Taken together, these findings suggest that perception of the whole person is influenced not only by the face, but also by the body form and motion. Nonetheless, previous studies did not provide a quantitative measure for the unique contribution of body form and motion beyond the face and the extent to which they contribute independently or interactively to the representation of the whole person.

Modeling representation presents a unique challenge, as it is unclear what is the most appropriate measurement for quantifying it. One can measure accuracy in recognition for

example, or the evaluation ratings of different traits (see for example Hu et al., 2018). Another possible measure that can be used for these purposes, and which we indeed decided to apply in the current study, are distinctiveness ratings. The main advantage of using distinctiveness ratings is that they involve comparing each identity to an average or a prototypical representation, and therefore are especially useful for modeling perception of person identity. Distinctiveness ratings have indeed been shown to be related to face recognition for example (i.e. Butcher & Lander, 2016) as well as the contribution of gait to recognition (Simhi & Yovel, 2019), indicating that they represent useful properties related to identity. These ratings are therefore a good measure which can be used to quantify the contribution of each of the different components in person perception (i.e. the face, body form and gait) to the whole person representation, as can be done using a general linear model, for example, which allows for the assessment of the contribution of each of the components, and whether it is independent or not.

In addition, since distinctiveness ratings involve comparison to a prototypical representation, they can also be used to map the different identities in the same perceptual space. This approach is complementary to linear modeling, and is usually based on the distances between identity pairs (i.e. their similarity/difference). Indeed, ratings of distinctiveness have been used in the face perception literature to generate a multi-dimensional face space (Lee, Byatt, & Rhodes, 2000; Valentine, 1991), in which the face space distance corresponds to perceptual similarity and the dimensions to different facial features (e.g. Abudarham & Yovel, 2016). Perceptual space based assessment methods have been used to account for many known phenomena in face recognition (Lee et al., 2000; Leopold, Bondar, & Giese, 2006; Rhodes & Jeffery, 2006; Valentine, 1991, 2001; Valentine, Lewis, & Hills, 2016), and findings suggest that

body encoding abides by the same principles (Rhodes, Jeffery, Boeing, & Calder, 2013; Streuber et al., 2016). Recent studies have used conceptually similar approaches in innovative ways in order to map language representational spaces to a body representational space (Hill et al., 2016), and assess how similarity in the verbal description of a body relates to similarity in body shape features.

In the current study, we therefore used two complementary modeling procedures, multiple regression and a multi-dimensional representational space analysis to assess whether the face, body and motion, each makes a unique contribution to the representation of the whole person in an independent or an interactive manner. To that end, we asked participants to rate the distinctiveness of identities in dynamic and static whole person stimuli based on their face, body form, gait or based on the person as a whole (see Figure 1 for a schematic display of the design). Using these ratings, we performed two different analyses: First, we used the distinctiveness ratings of the components (face, body form and gait) as predictors in a linear regression model of the distinctiveness ratings of the whole person. We modeled static and dynamic whole person perception separately, adding each component of the whole person to the model one at a time in order to assess its contribution to the representation. We start by modeling whole person perception based on the face alone, as the default model, and then assess if body form contributes to the prediction of whole person ratings beyond the face. Finally, in the case of dynamic person perception, we also examine if gait contributes beyond the face and body form. This method enabled us to examine the unique contribution of each component, and also to examine whether their contributions are independent or interactive.

Second, to complement the linear regression analysis, we conducted an additional analysis in which we used the different components of the whole person as dimensions, and mapped the different identities to locations in a multi-dimensional perceptual space of the identity of the whole person. In order to examine whether the face, body form and motion are valid dimensions of the multi-dimensional representational space of the whole person, we computed the Euclidian distances between all pairs of identities in the multi-dimensional perceptual space based on ratings of these different components, adding one component at a time to create different representational spaces. We then measured the correlations between these Euclidian distances and the differences between the actual whole person ratings in order to assess if each additional dimension in the representational space (i.e. – each additional component) contributed to the whole person representation (see Figure 2 for a schematic presentation of this procedure). In this analysis also, we first assessed the contribution of the face alone, as the default model, and then assessed whether the body and gait contribute to person perception beyond the face.

Methods

Participants. 114 participants took part in this experiment, in exchange for either course credit or payment. Two participants were excluded from analysis – one who could not complete the experiment due to a technical error, and one who did not comply with the experimental instructions. Thus, 112 participants were included in the final experiment (91 female, mean age = 23.03, SD = 3.68), 16 in each of seven experimental rating conditions (static: whole person, face and body rating; dynamic: whole person, face, body form and gait rating). A similar sample size was used in our previous studies of whole person perception (Simhi & Yovel, 2016, 2017). The

analysis, however, is not based on the variance between participants but on the variance between their averaged ratings of the different stimuli and therefore power calculations were performed to determine the required number of stimuli in the experiment, rather than the number of participants as typically done. We address this in the Stimuli section below. All participants were recruited at Tel Aviv University, had normal or corrected-to-normal vision and gave their informed consent to participate in the study by signing the appropriate consent form approved by the Tel Aviv University ethics committee.

Stimuli. The stimuli in the experiment consisted of 89 Caucasian identities (66 female) that were selected from the Database of Moving Faces and People (O'Toole et al., 2005). For each identity, two different stimuli were included in the experiment, filmed on different occasions (up to six months apart) making a total of 178 stimuli. Analysis was conducted across stimuli and therefore the power of the analysis is determined by the number of the stimuli. We estimate the minimum number of stimuli required in the following manner: Assuming a medium effect size of .15, alpha of .05, and power of .8 we examine the required number of stimuli in order to create a model with an R^2 significantly different from zero using three predictors (which is the largest number of predictors in the planned models in this study, as detailed in the Data Analysis section below). 77 stimuli are required in this case, therefore using 89 identities in our experiment is expected to yield sufficient power. Furthermore, in an attempt to replicate our findings within the same study, we repeated the same experiment twice within-participant: We first collected rating data from the first set of 89 stimuli and then collected the same rating data using the second, different set of stimuli, of the same 89 identities. We conducted two separate analyses for each set of stimuli to assure the results are replicated.

Stimuli were displayed either as dynamic or static in different experimental conditions. Dynamic stimuli consisted of 5 seconds of video, depicting the identity walking towards the observer in a naturally lit corridor. Static stimuli consisted of one of the final frames of the aforementioned videos, which was selected so that the identity was in as static a position as possible (and not mid-stride for example). This image was presented for 5 seconds in the static display conditions.

It should be noted that the relatively large number of stimuli presented in this study (89 identities X 2 filming occasions = 179), was chosen specifically to encourage norm-based ratings. Nonetheless, even if relative ratings were used, the relatively large selection, taken from a naturalistic dataset, attempts to create a situation where these ratings are as close to the norm as possible.

Both videos and static images were displayed at roughly 18.3° by 12.3° visual angle during the experiment and took up 720 × 480 pixels on the screen.

Design, Apparatus and Procedure. The experiment was presented in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on a Samsung SyncMaster SA950, Full HD, LED monitor with a 1920 × 1080 screen resolution, in front of which the participants were seated at a comfortable distance of approximately 60 cm.

Participants were asked to rate the distinctiveness of either the face, body form or whole person from static images of the different identities, or the distinctiveness of the face, body form, gait or whole person from dynamic displays of these different identities. Thus, the study included a total of 7 different, between participant, rating tasks. For each of the 7 conditions, participants

were presented first with 89 stimuli depicting each of the 89 identities in the experiment and then with another 89 stimuli depicting the same identities, in a different set of stimuli (filmed on a different occasion). Within each 89 stimulus stretch, the order of presentation was randomized. In all rating conditions the static and dynamic stimuli depicted the whole person, however participants in the different tasks were asked to focus on one of the components based on the condition they were assigned to. See Figure 1 for a schematic design of this procedure.

For each stimulus in the experiment, participants were asked to rate the distinctiveness of the identity in the stimulus on a scale of 1-7. In order to explain the concept of distinctiveness to the participants, in each condition they were asked to estimate (for face ratings as an example): “How unique and different is this person’s face as compared to that of other people – e.g. how easy do you think it would be to recognize this person based on their face as compared to recognizing other people? A rating of 7 for example would indicate that the person’s face is very unique and different and that it would be easy to differentiate between his face and other people’s faces. A rating of 1 on the other hand would indicate that the person’s face is not very unique and different and that it would not be easy to differentiate between his face and other people’s faces.” Face in this example was replaced by the component being rated in each case. These instructions were given in Hebrew.

After presenting the instructions and ensuring that the participants understood them, the experiment began. On each trial a stimulus was presented for 5 seconds. After the stimulus presentation a rating scale appeared and participants were asked to rate the component they were assigned to on a scale of 1-7 (see Figure 1). After making their response a new trial began.

Unbeknownst to the participants, the first 89 trials depicted 89 different identities, which appeared again in the second half of the experiment, in a different set of stimuli, which were filmed on a different occasion.

The experiment included 10 second breaks every 20 trials, with 178 trials in total.

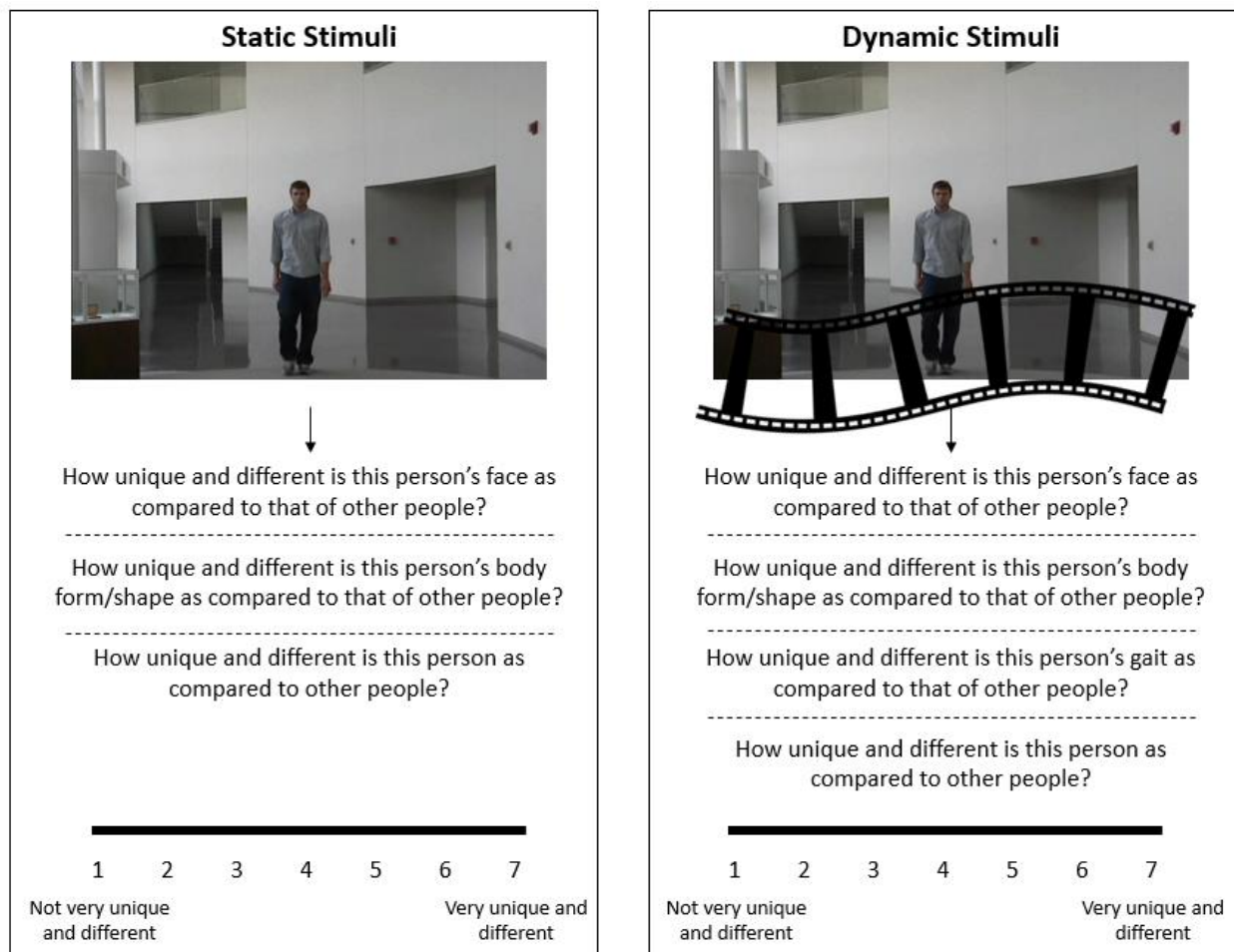


Figure 1. Experimental design and procedure: The experiment included a total of seven between-participant conditions. The rating conditions were either face, body form and whole person for the static stimuli conditions (left panel), or the face, body form, whole person or gait in the dynamic stimuli conditions (right panel). Participants in each group viewed a stimulus of the whole person for 5 seconds and afterwards were asked to rate the distinctiveness of the identity in the stimulus based on the condition they were assigned to.

Data Analysis. We removed from analysis all trials with response times (RTs) longer than 10 seconds, and scaled the ratings of each participant to a scale of 1-7. Next, we averaged the responses across participants, per identity (89), stimulus presentation (2 – each identity was presented twice, in different stimuli, as described above) and presentation rating condition (7 - static/dynamic face, body, whole person or gait). The two presentations of each identity were used in order to replicate the results within the same study. As detailed in the Stimuli section, these two presentations consisted of different videos/images of the same identity, filmed on different days. Subsequent analyses were thus conducted based on two sets of 623 (89 X 7) averaged, stimulus-based responses.

Rating reliability

To examine rating reliability, we first examined the distributions of the ratings in each of the conditions, in each of the two stimulus presentations. This allowed us to visually assess if the distribution of ratings was similar across the first and second presentation of the different identities. Next, we examined the Pearson correlations between the ratings of the first and second stimulus of each identity, and compared them to the correlation between random pairings of identities in the first and second stimulus presentations.

Regression analysis

The ratings of the different components of the whole person were used as predictors for dynamic and static whole person ratings in a linear regression model. We first report the zero-order correlations between the different components and the whole person, and then perform the multiple regression analysis, for each of the two presentations of each identity. We

performed this analysis for dynamic and static whole person representations separately, based on the hypothesized models: since most of the studies on person perception rely on stimuli which include the face alone, we start with a model which includes only the face, and then examine if the addition of body form better predicts whole person ratings, thus, contributing to the representation beyond the face. Finally, in the case of dynamic representation, we assess if gait contributes to prediction beyond the face and body form. In this way we tested the unique contribution of each of the components to the representations of the whole person, in dynamic and static conditions. For the static model only ratings of static stimuli were used, and for the dynamic model only ratings of dynamic stimuli were used. We also examined the possible contribution of paired and triple interactions between the components in each model. While all regression analyses are reported without mean-centering of the variables, we examined the regression models with mean-centering as well, and revealed similar findings.

A perceptual space of the whole person

Following the regression analysis, we conducted an additional, unplanned analysis that examined whether the components of the whole person are valid dimensions of its representational space. To examine whether the face, body and gait are valid dimensions of the representational space of the whole person, we calculated the absolute values of the difference between the ratings of the whole person for each pair of identities, for each viewing occasions (i.e the difference between the ratings in each of the 89 stimuli and each of the remaining 88 stimuli: $89 * 88 = 7,832$; since the absolute value of the difference is symmetrical this left us with $7,832/2 = 3,916$ possible pairings in total), in the dynamic and static conditions. We then computed the Euclidean distances in a multi-dimensional space of 1 to 3 dimensions, which could

include the face, body form and gait (for dynamic stimulus representations). As before, 3,916 such distances were computed in each multi-dimensional space, for each stimulus presentation, representing the distances between each pair of identities. Finally, we calculated the Pearson correlation between the differences in whole person ratings and the Euclidian distances calculated based on the representational space created using the different combinations of whole person components. These correlations indicated whether these components are valid dimensions of the whole person representation. See Figure 2 for a schematic representation of this analysis. We performed this analysis for dynamic and static whole person representations separately, based on the hypothesized models detailed in the linear regression section, and in the same order. As in the linear regression analysis, for the static representational space only ratings of static stimuli were used, and for the dynamic representational space only ratings of dynamic stimuli were used.

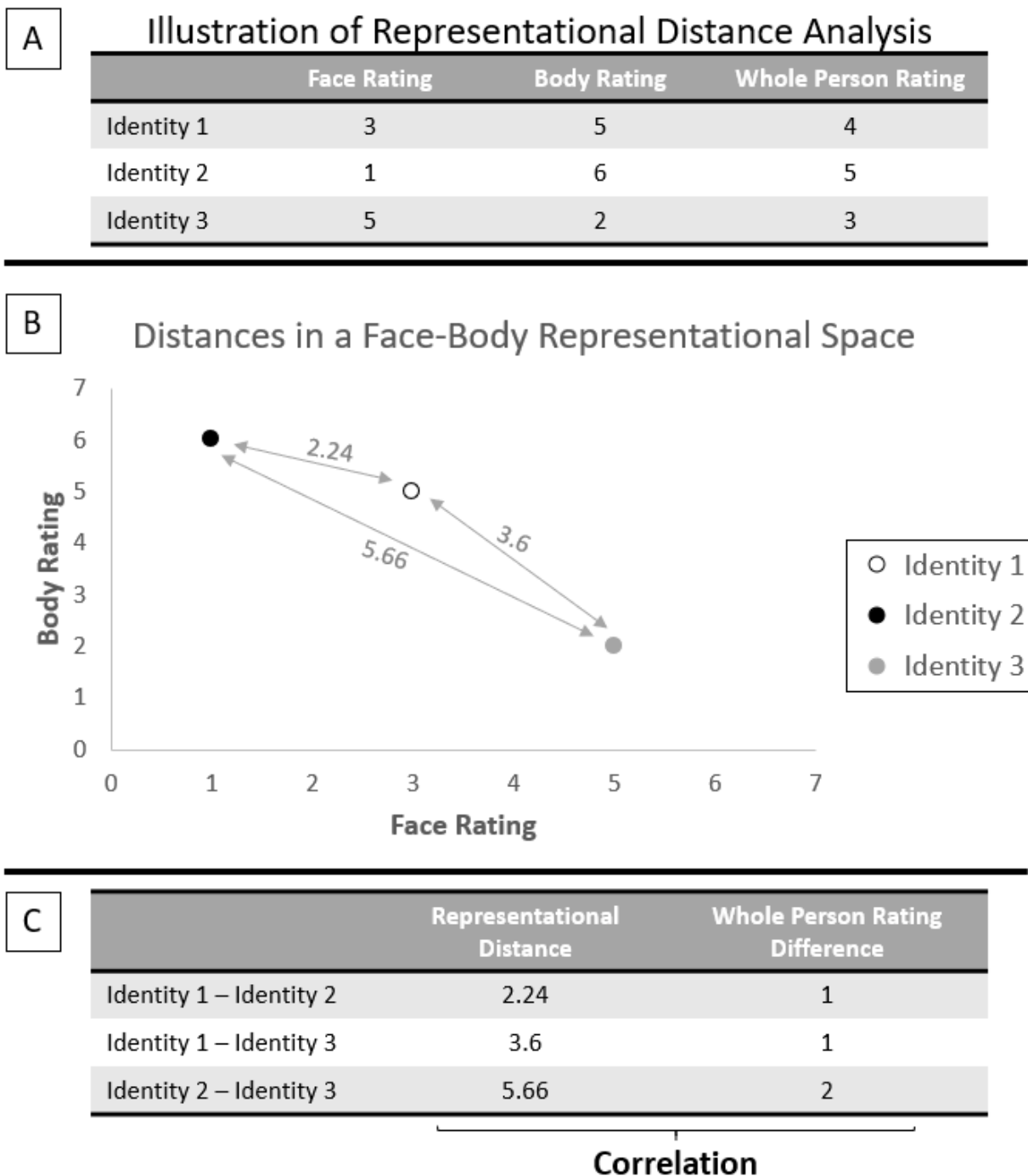


Figure 2. To demonstrate how representational distances were calculated we present data of 3 made-up identities. Panel A shows the distinctiveness ratings of the face, body and whole person for each of the 3 identities. Panel B shows the calculation of the Euclidean distances between the identities, based on the face and body ratings, in a face-body representational space. Panel C shows the representational distances, calculated using the representational space, and the actual differences (in absolute values) between the whole person ratings of the identities (shown in Panel A). The correlation was calculated between these two measures, across the different

identities. This method was used to assess the unique contribution of the face and body to the whole static person and the unique contribution of the face, body and gait to the whole dynamic person.

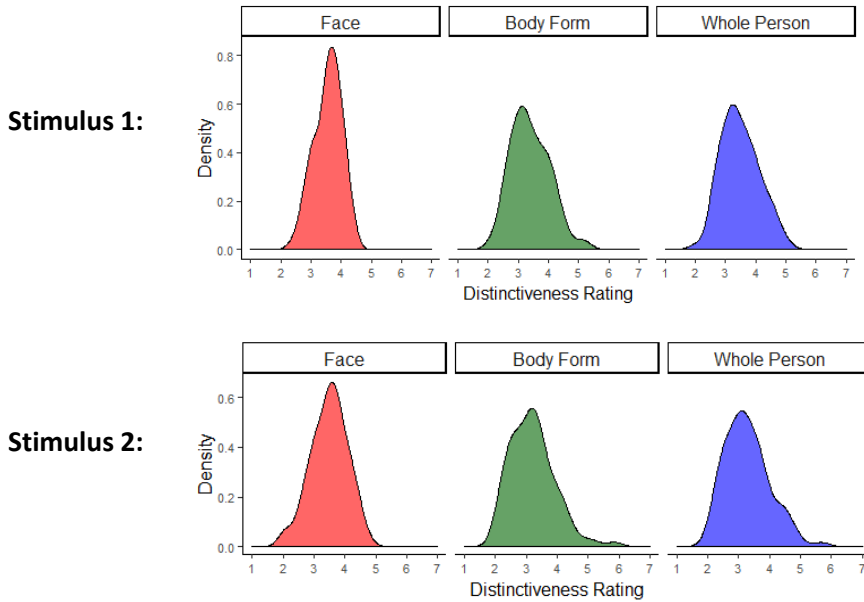
Data processing was performed using Python. Statistical analysis was performed in JASP (Version 0.10.0) and RStudio, with comparison between correlations being conducted using the `cocor` package (Diedenhofen & Musch, 2015) using Zou's confidence interval (Zou, 2007). Multiple comparisons were corrected by using the Bonferroni method. Sample size analysis was performed using GPower 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007).

Results

Rating distribution and consistency

To assess the reliability of distinctiveness ratings, we examined how consistent were ratings of the same identity when displayed in different stimuli. To do so we first examined if the distribution of the ratings was similar in the first and second stimulus presentations of each identity (Figure 3). Next, we examined the correlations between the ratings of the first and second stimulus of each identity, in each of the experimental conditions, and compared these correlations to correlations between random pairings of first and second stimulus trials (Figure 4).

A. Static Stimuli:



B. Dynamic Stimuli:

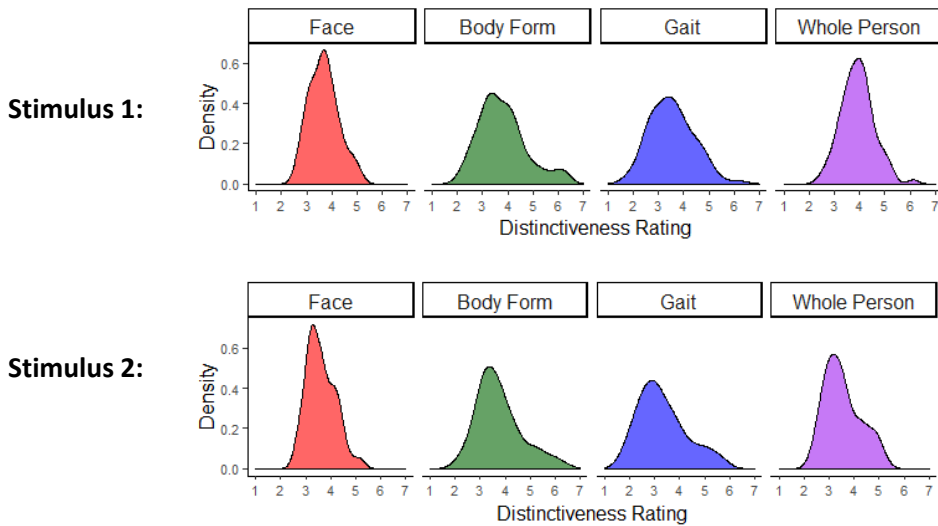


Figure 3. Density plots depicting the distribution ratings of the first and second stimulus of each identity, for each rating condition. Panel A depicts the rating for static stimuli and Panel B depicts the rating for dynamic stimuli.

Overall the distributions appeared relatively similar between the first and second presentation of each stimulus.

Next, we calculated the Pearson correlations between the ratings of the first and the second stimulus of the same identity, and the correlations between randomized pairings of identities presented in the first and second stimulus set, as can be seen in Figure 4. We found that for all ratings except the static face, in both static and dynamic displays the correlation between the same-identity pairings as compared to the randomized pairings was significantly higher, as assessed using Zou's two-sided confidence interval (Zou, 2007) for dependent, overlapping groups. The results of these comparisons can be seen in Table 1, after Bonferroni correction for seven comparisons. Overall these findings indicate that in nearly all cases, the ratings of the whole person and its components were reliable, and suggest that distinctiveness is relatively consistent between different viewings of the same identity.

As a complementary analysis, we also computed the correlations between the two presentations of the same identity within each participant and averaged these correlations across participants. This analysis revealed correlations that were weaker (r (range) = .15 – .41) but were also all significantly higher than random pairings.

Next, we assessed whether the face, body and gait are valid predictors of the whole person.

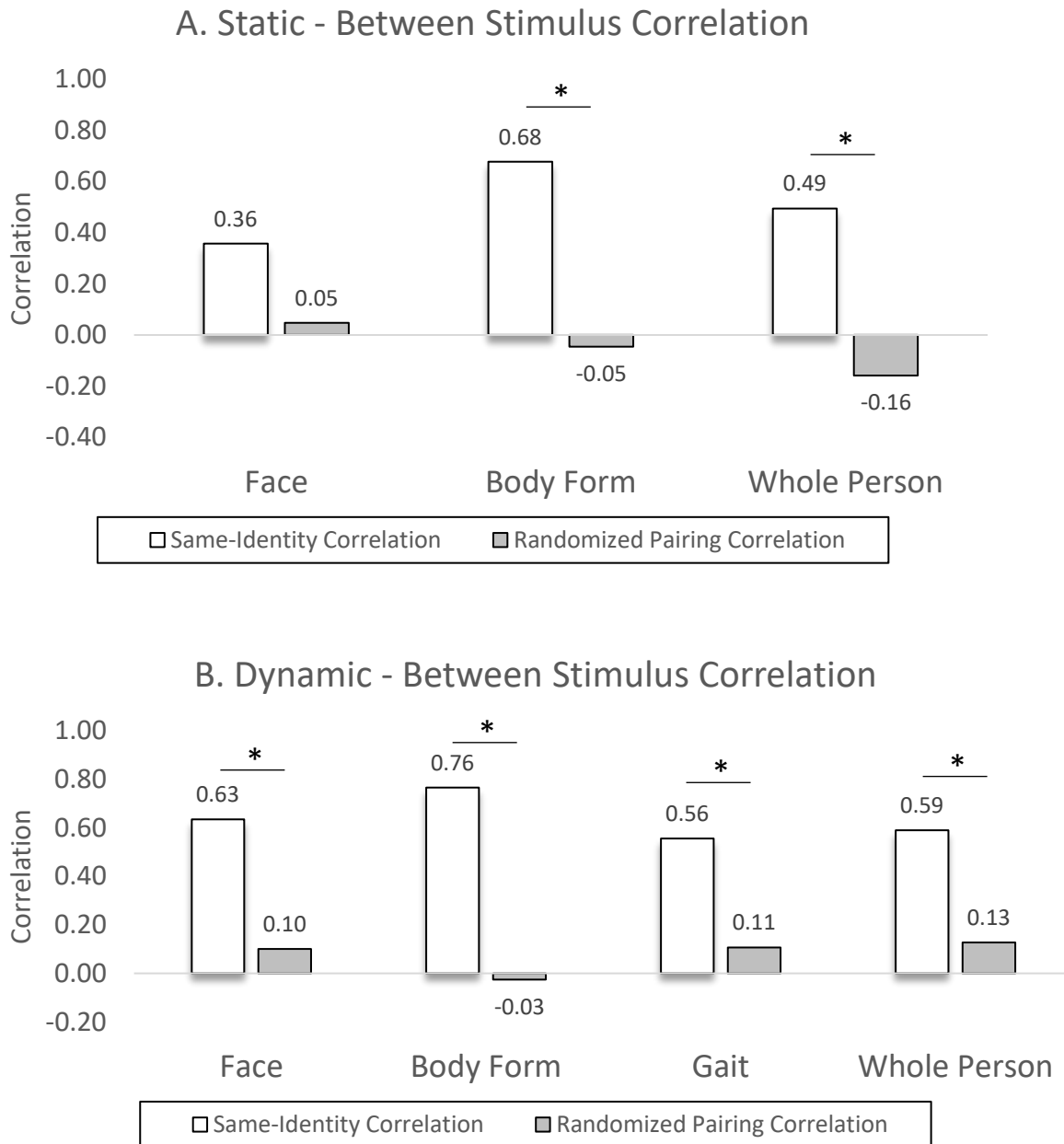


Figure 4. The Pearson correlations between the ratings of the first and second stimulus of each identity, for each rating condition in the experiment, appear in white and the correlations between random pairings of stimuli appear in gray. Except for static face rating, all same-identity correlations were significantly higher than the randomized pairing correlations. Panel A depicts the correlations for static stimuli, and panel B depicts the correlations for dynamic stimuli. Significant differences after Bonferroni correction for seven comparisons ($p < .007$), are marked with *.

Table 1

The Difference between the Correlations of Same vs. Random Identity Pairings.

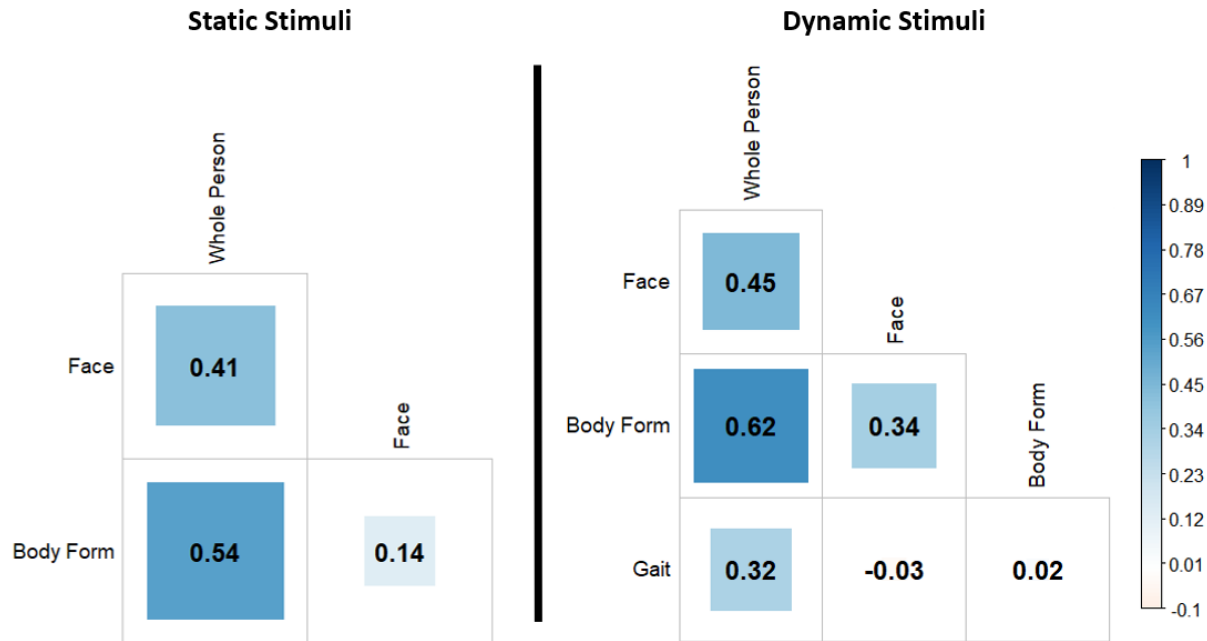
Measure	Static			Dynamic			
	Face	Body form	Whole person	Face	Body form	Gait	Whole person
Correlation difference (same – random pairing)	.31	.73	.65	.53	.79	.45	.46
99.3% CI	-.10-.69	.38-1.04	.30-.95	.20-.85	.48-1.08	.10-.78	.15-.76

Note. 99.3% confidence intervals are presented, after applying Bonferroni correction for seven comparisons.

Regression analysis

We first report the zero-order correlations between the different components used in the regression analysis, and for each of them with the whole person representation, in the static and dynamic conditions. The correlations are shown in Figure 5, for each of the stimulus sets. Overall, we see that the correlations between the different components were lower than their correlations with the whole person. This enables us to estimate their unique contributions to whole person representation.

Stimulus 1:



Stimulus 2:

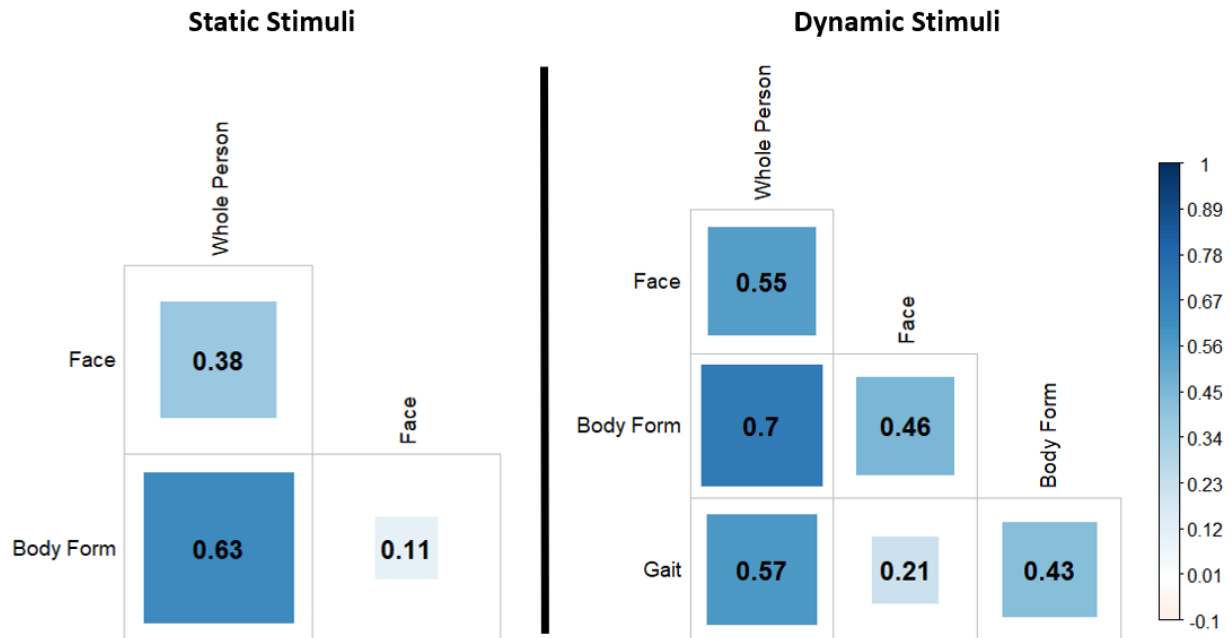


Figure 5. Pearson correlations between the rating conditions: The Pearson correlations between the ratings of the whole person and its components for the dynamic and static stimuli, in each of the two stimuli sets.

Next, to examine the contribution of different components to whole person representation, we performed linear regression separately for the prediction of static and dynamic whole person ratings, for each stimulus set. We started by examining a model based on the face alone, as is most common in person perception studies, and then assessed the possible contribution of body form in the static and dynamic whole person model, and also the gait in the dynamic whole person model. Finally, we also assess possible interactions between the different components and their contribution to the whole person representation.

Static whole person: this analysis is based on face, body form and whole person ratings of the static whole person. In order to estimate the contribution of each component to the representation of the static whole person, we created two linear regression models, for each stimulus set. We first examined a model of whole person ratings based on the ratings of the face alone. Next, we examined a model including both the face and body form. In both of these models, in both stimulus sets, all of the predictors were significant at $p < .001$. See Table 2 for a full summary of the adjusted R^2 , standardized betas and significance values of each component in the models. Figure 6 depicts a comparison between the model predictions and actual whole person ratings based on the face rating alone and the combination of the face and body ratings.

Table 2

Model Comparison for Prediction of Static Whole Person Ratings

Stimulus Presentation	Components included	R^2_{adj}	Face			Body form		
			beta (standardized)	t	p-value	beta (standardized)	t	p-value
First	Face	.16	.41	4.23	<.001	-	-	-
	Face + Body form	.40	.34	4.12	<.001	.50	5.95	<.001
Second	Face	.13	.38	3.78	<.001	-	-	-
	Face + Body form	.48	.31	4.03	<.001	.60	7.78	<.001

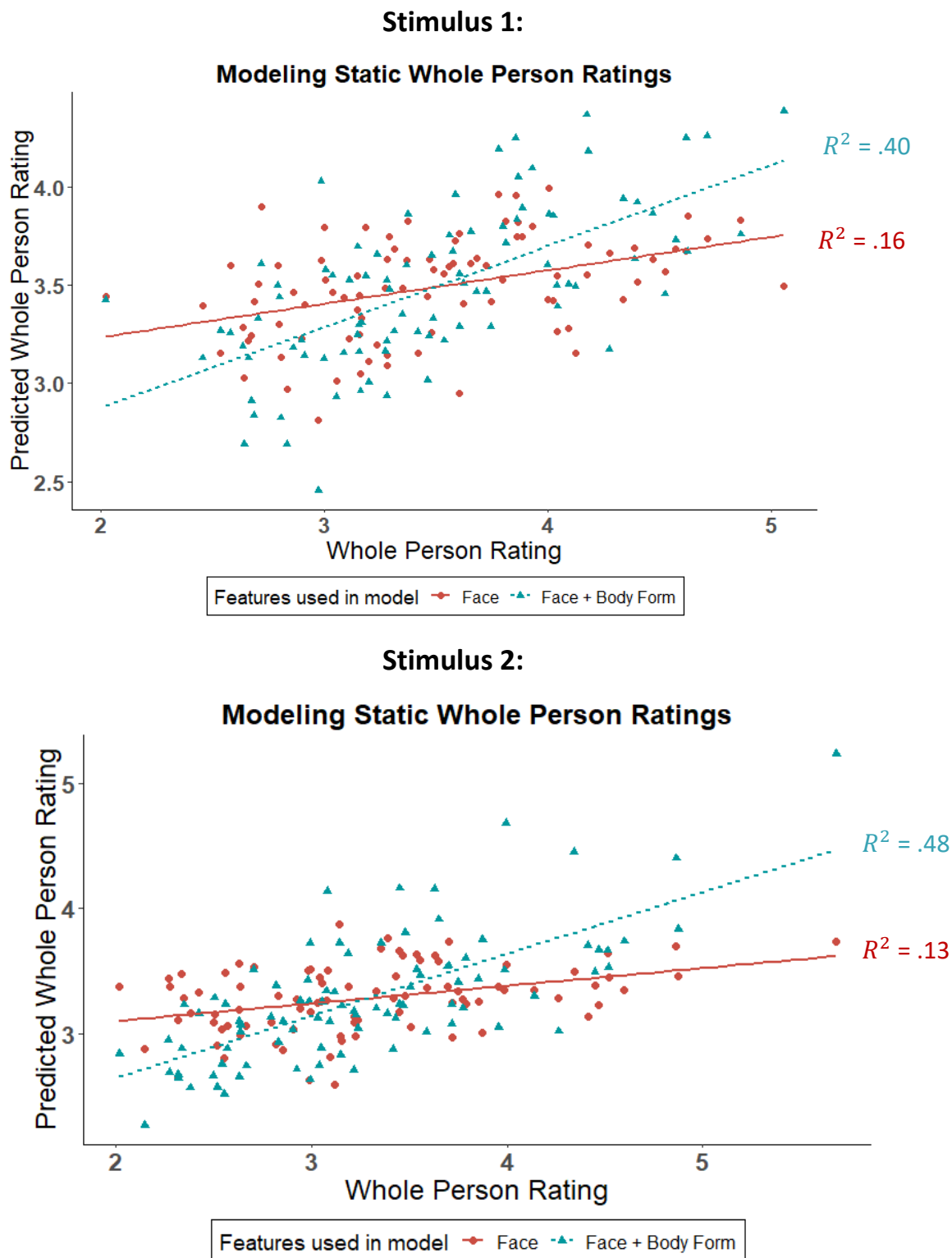


Figure 6. Predicted ratings based on the face or the face and the body vs. actual ratings in models of static whole person perception, for each of the two sets of stimuli of each identity.

Model comparison revealed that indeed the model using both the face and body form to predict static whole person ratings, was better than the model including the face alone in both the first presentation of each identity ($F(1,86)=35.38, p<.001$) as well as the second ($F(1,86)=60.57, p<.001$). These results indicate that the face and body form both contribute to the representation of the static whole person.

Dynamic whole person: this analysis is based on face, body form, gait and whole person ratings of the dynamic whole person. In order to estimate the contribution of each component to the representation of the dynamic whole person, we created three linear regression models, for each stimulus set. First, we modeled the dynamic whole person based on the face alone. Next, we added the body form to the model, and finally, we added the gait. In all three of these models, in both stimulus sets, all of the predictors were significant at $p<.005$. See Table 3 for a full summary of the adjusted R^2 , standardized betas and significance values of each component in each of the models. Figure 7 depicts a comparison between the model predictions and actual whole person ratings.

Table 3

Model Comparison for Prediction of Dynamic Whole Person Ratings

Stimulus Presentation	Components included	R^2_{adj}	Face			Body Form			Gait		
			beta (standardized)	t	p-value	beta (standardized)	t	p-value	beta (standardized)	t	p-value
First	Face	.19	.45	4.66	<.001	-	-	-	-	-	-
	Face + Body form	.43	.27	3.16	.002	.53	6.20	<.001	-	-	-
	Face + Body form + Gait	.53	.28	3.62	<.001	.52	6.66	<.001	.32	4.30	<.001
Second	Face	.30	.55	6.19	<.001	-	-	-	-	-	-
	Face + Body form	.55	.29	3.67	<.001	.57	7.09	<.001	-	-	-
	Face + Body form + Gait	.64	.29	3.97	<.001	.43	5.54	<.001	.33	4.58	<.001

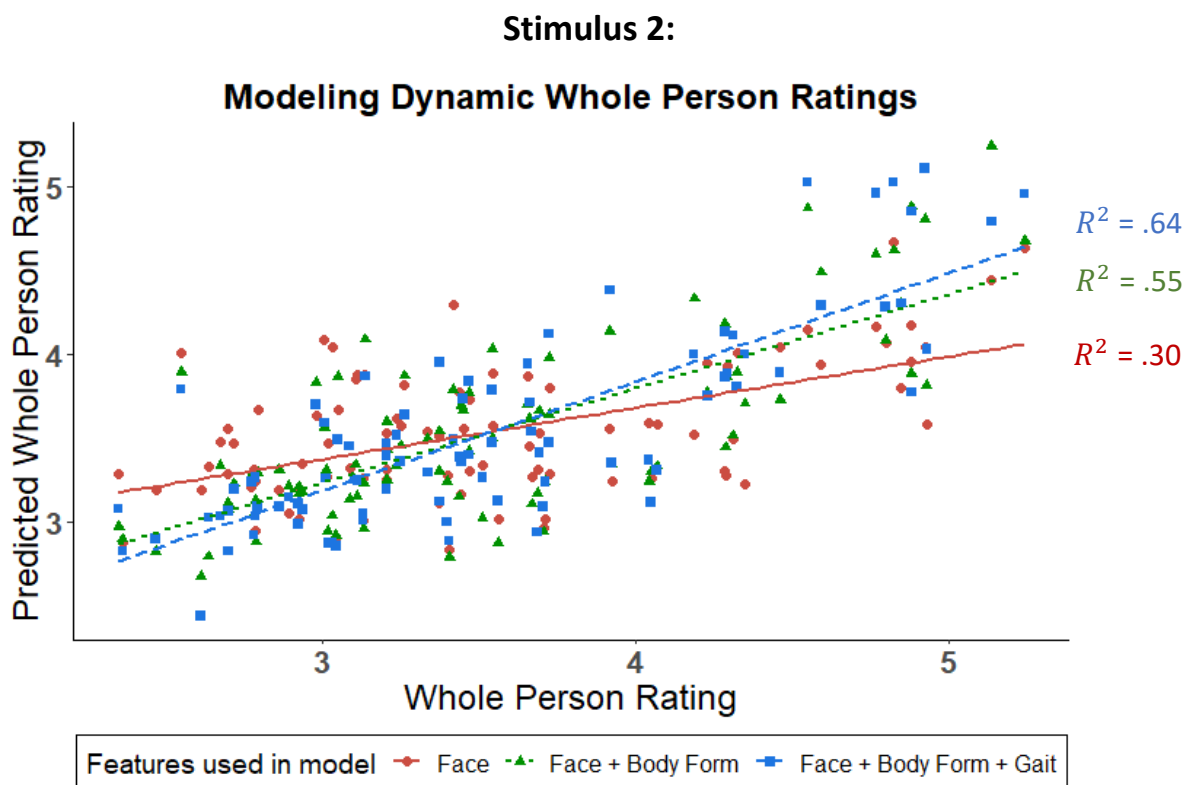
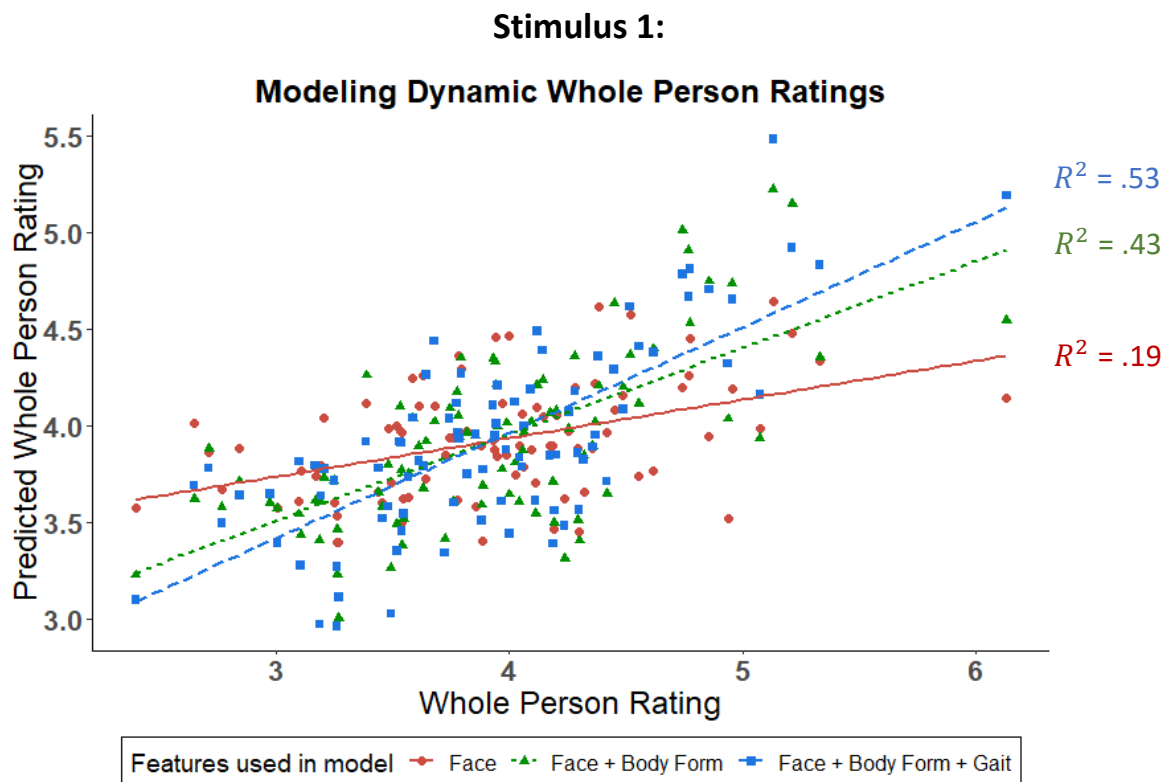


Figure 7. Predicted rating based on the face, face and body and face, body and gait vs. actual ratings for dynamic whole person perception, for each of the two sets of stimuli of each identity.

Model comparison revealed that the model based on the dynamic face and dynamic body form was significantly better at predicting dynamic whole person ratings compared to the dynamic face alone in both the first ($F(1,86)=38.49, p<.001$) and second stimulus of each identity ($F(1,86)=50.3, p<.001$). Furthermore, the model based on the dynamic face, body form, and gait was better than the model based on the dynamic face and body form alone in both the first ($F(1,85)=18.51, p<.001$) and second stimulus set as well ($F(1,85)=20.96, p<.001$). These results indicate that the face, body and gait all have unique contributions to the representation of the dynamic whole person.

We next examined whether adding interactions between the components in our models will provide a better fit to the whole person representation. To do so, we examined each possible paired interaction in the static and dynamic models (static model: face X body form; dynamic model: face X body form, face X gait, body form X gait) and the triple interaction in the dynamic model (face X body form X gait). In both static and dynamic models, in both stimulus sets, model comparison revealed that the interactions did not contribute beyond the best model outlined in the previous section. These results are outlined in Table 4 for the static model and

Table 5 for the dynamic model.

Table 4

The Interactions Between the Whole Person Components in Static Whole Person Perception

Modeling

Stimulus Presentation	Components included	R^2_{adj}	Comparison to the best model without interaction
First	Face + Body form	.40	-
	Face + Body form + Face X Body form	.39	$F(1,85)<1, p=.93$
Second	Face + Body form	.48	-
	Face + Body form + Face X Body form	.48	$F(1,85)<1, p=.92$

Note. The table includes the R_{adj}^2 of each of the examined model with interactions, as well as the optimal model without interactions (highlighted in light gray), and the results of the statistical comparison between these models, for both sets of stimuli.

Table 5

The Interactions Between the Whole Person Components in Dynamic Whole Person Perception

Modeling

Stimulus Presentation	Components included	R_{adj}^2	Comparison to the best model without interaction
First stimulus	Face + Body Form + Gait	.53	-
	Face + Body form + Gait + Face X Body form	.52	F(1,84)<1, p=.92
	Face + Body form + Gait + Face X Gait	.52	F(1,84)<1, p=.68
	Face + Body form + Gait + Body form X Gait	.53	F(1,84)=1.34, p=.25
	Face + Body form + Gait + Face X Body form X Gait	.53	F(1,84)<1, p=.55
Second stimulus	Face + Body Form + Gait	.64	-
	Face + Body form + Gait + Face X Body form	.63	F(1,84)<1, p=.70
	Face + Body form + Gait + Face X Gait	.63	F(1,84)<1, p=.96
	Face + Body form + Gait + Body form X Gait	.63	F(1,84)<1, p=.39
	Face + Body form + Gait + Face X Body form X Gait	.63	F(1,84)<1, p=.68

Note. The table includes the R_{adj}^2 of each of the examined models with interactions, as well as the optimal model without interactions (highlighted in light gray), and the results of the statistical comparison of each of these models to the optimal model, for both sets of stimuli.

Representational Space Analysis

In order to assess the whole person representational space, we created a multi-dimensional perceptual space and assessed whether the face, body form and gait are valid dimensions of this space. In order to do so we calculated the distances between each possible

pair of identities in the experiment in this representational space using different components as dimensions in the space. Then, we calculated the correlation between these distances and the differences between the actual whole person ratings of these identities, separately for dynamic and static conditions, and for the two stimuli sets. See Figure 2 for a schematic presentation of this procedure and the Methods section for details. In this way, we compared the representation based on whole person ratings with that of a representational space based on combinations of different whole person components.

Static whole person: this analysis is based on face, body and whole person ratings of the static whole person. Results show that the distances between the rated components were best correlated with the differences in static whole person ratings when both the face and body form were included as dimensions in the representational space: the correlation increased significantly when body form was added to the representational space as an additional dimension, on top of the face, as assessed using Zou's two-sided confidence interval (Zou, 2007) for dependent, overlapping groups (First stimulus set – $r_{\text{face+body form}} - r_{\text{face}} = .22$, 95% CI [.19 .25]; Second stimulus set – $r_{\text{face+body form}} - r_{\text{face}} = .30$, 95% CI [.27 .33]).

Dynamic whole person: this analysis is based on face, body form, gait and whole person ratings of the dynamic whole person. The highest correlations were found when the face, body form and gait were all included as dimensions in the representational space. The correlation increased significantly when body form was added as a dimension on top of the face (First stimulus set – $r_{\text{face+body form}} - r_{\text{face}} = .20$, 95% CI [.17 .23]; Second stimulus set – $r_{\text{face+body form}} - r_{\text{face}} = .17$, 95% CI [.14 .19]), and when gait was added as a dimension on top of the face and body form

(First stimulus set – $r_{\text{face+body form+gait}} - r_{\text{face+body form}} = .05$, 95% CI [.03 .07]; Second stimulus set – $r_{\text{face+body form+gait}} - r_{\text{face+body form}} = .02$, 95% CI [.003 .04]).

The correlations of the different models are plotted in Figure 8.

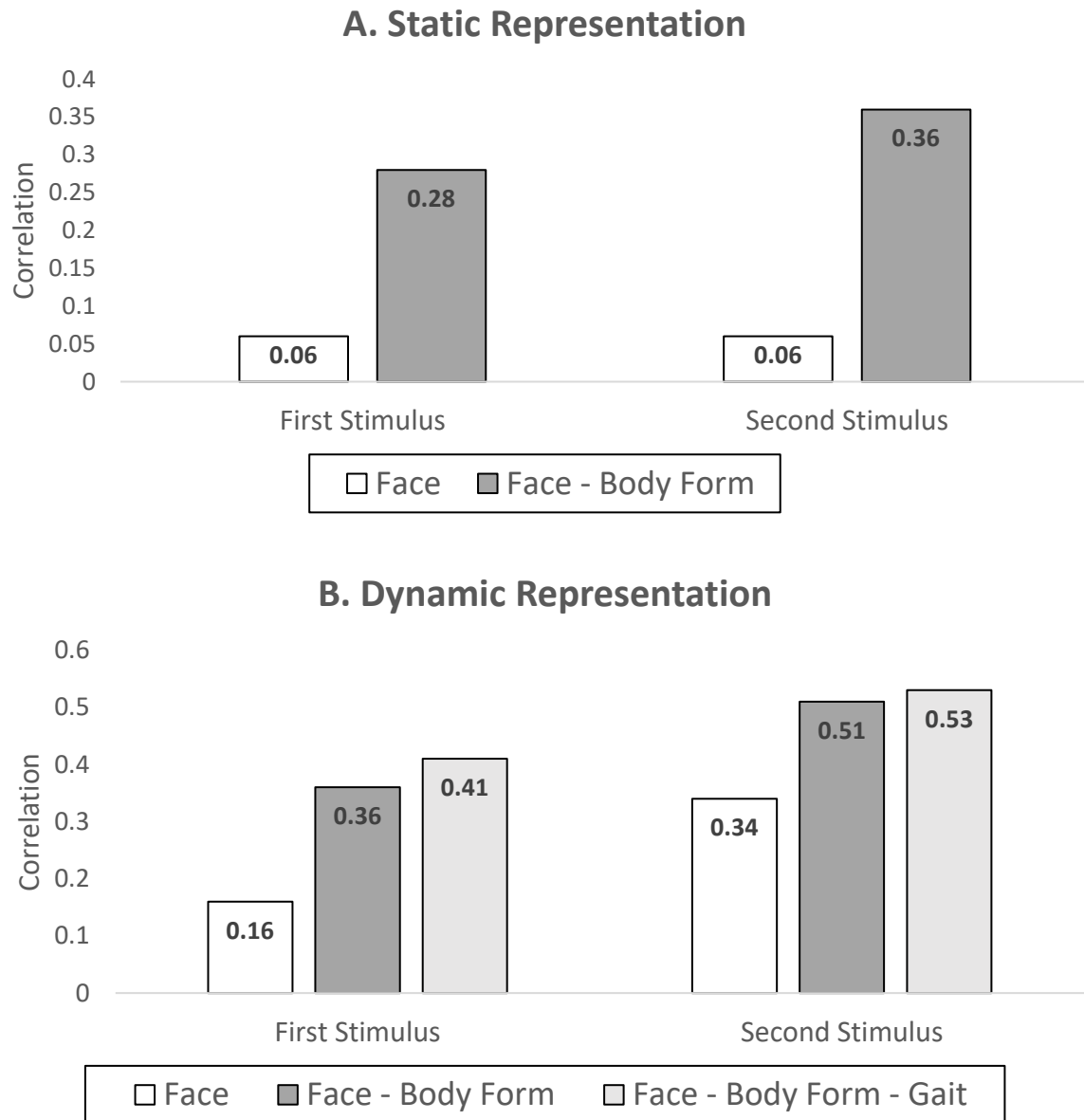


Figure 8. Whole person representational space analysis: The correlations between the distances between identities based on whole person ratings, and based on the Euclidian distances in a representational space using different components as dimensions. A – static whole person correlations. B – dynamic whole person correlations. Correlations for both stimulus sets are depicted.

Discussion

The goal of this study was to assess whether the face alone can account for the representation of the whole person, or whether the body and gait may contribute beyond the face and better account for whole person representation. This is the first time that the unique contribution of these different components was directly assessed in a quantitative manner, for both static and dynamic whole person perception. The results clearly show that the representation of the whole person is determined by the face and the body form, and in the case of a dynamic person, also the gait. Each of these different components contributed independently and significantly to the representation of the whole person. This result was found in two complementary analyses that showed that the representation of the person is best predicted by a combination of its components, the face and body form for static images and also the gait for dynamic stimuli. First, a linear regression analysis demonstrated a contribution of the body beyond the face, with an additional contribution of gait in the case of dynamic person representation. Next, an additional representational space analysis, based on distances between the ratings of different identities, again showed that the representational space that is based on both the face and the body form provided a better fit to the representational space of the whole person than when using the face alone. In dynamic person perception, adding gait as the third dimension, in addition to the face and body, provided a better fit to the dynamic whole person representational space as compared to the face and body alone. These findings were replicated in two different stimulus sets of each identity, using stimuli while were filmed on different days. Together, these analyses indicate that person perception goes beyond the face and that the body and gait also contribute uniquely and significantly to the representation of the whole person.

The regression model was also used to assess the contribution of possible interactions between face, body form and gait perception. We found that in both static and dynamic whole person representation, interactions between the face and body did not contribute to whole person representation. No significant interactions were found with gait as well, in dynamic whole person perception. This indicates that not only do the face, body and gait contribute to whole person representation, but their contribution is independent.

An additional finding in this study was that the representation of the same identity is relatively stable over time. When comparing the ratings of the same identity in stimuli which were filmed on different occasions (up to six months apart), we found that there was a higher correlation between the ratings of the same identity on different occasions (in the different stimuli) as compared to randomized pairings between the identities, in all of the rating conditions except for the static face. We presume that consistency ratings were lower for faces because faces are relatively small and less salient in whole person stimuli. Performing this study while having later frames of the video available, showing the faces in higher resolution, is highly likely to increase the consistency of static face ratings, and having additional static information available by using displays of multiple static images, as opposed to a single static image, is likely to increase consistency as well. The high consistency of ratings of the body and gait is especially interesting since the visible shape of the body is significantly affected by the clothes a person is wearing. This finding indicates that the main features of the body, which may make it distinct or non-distinct, are likely to be visible on different days and with different sets of clothes. In regards to gait, while analysis of gait is sometimes used in forensic situations today (see Larsen et al., 2008), indicating that the judicial system treats gait as a stable identifying feature, there are

relatively few studies examining this directly, to our knowledge. We have previously demonstrated that dynamic identity signatures in gait, i.e. idiosyncratic motion patterns of an individual (see O'Toole, Roark, & Abdi, 2002 for a description of these in the context of facial motion), contribute to familiar person recognition (Simhi & Yovel, 2017) and even unfamiliar person recognition (Simhi & Yovel, 2019), and recent findings suggest that they may be especially important when performing person recognition from a distance (see Hahn et al., 2015, Simhi & Yovel, 2019 and Yovel & O'Toole, 2016). The finding in the current study, that gait perception is relatively consistent between stimulus presentations in naturalistic videos that were filmed several weeks apart, supports these findings and indicates that gait may indeed be informative and useful for person perception.

It is important to note that in all of the conditions in the current study the whole person was always visible. Therefore, the ratings gathered for the face, body and gait may have influenced each other (for example, the face may have influenced the body rating). Indeed, these ratings are somewhat correlated, as can be seen in Figure 5. In addition, the mode of display, i.e. static vs. dynamic, is likely to have influenced these ratings as well, since, as can be seen in Figure 5, face and body form perception were correlated with gait perception in some of the stimulus presentations, indicating that participants were most likely basing their decision on some combination of static and dynamic parameters. We intentionally displayed the whole person in all cases here since we were interested in examining the independent contribution of each of these components, when all of the information is available. Nonetheless, it would be interesting to further examine the evaluations of these components in isolation, for example, the rating of the face when the body is not in view, and assess how this influences the results presented here.

Such an assessment would indicate how these components influence perception in isolation, when no mutual influences occur.

In conclusion, these findings stress that studies of person perception should go beyond the face and examine the contribution of the body and gait. In daily life we perceive the whole person, most often in motion, and our findings clearly demonstrate that real-life person perception cannot be reduced to perception of the face alone. Furthermore, the perceptual judgment which was examined in this study, distinctiveness, proved to be stable for each identity, across stimuli which were filmed on different days, in nearly all conditions, and results of regression and representational space analyses were replicated across the two stimulus sets.

These results can be viewed as support for the framework of whole person representation presented in Yovel & O'Toole, 2016. This framework suggests that the face, body form and gait are all important sources of information for the representation of the whole person. During person recognition for example, information from the face is used when recognizing people from a short distance, while information from body form and gait is important for person recognition from a distance. While distance was not manipulated in the current study, recent studies manipulating distance indeed support this idea, and demonstrate that the body is mainly used when recognizing people at a distance (Hahn et al., 2015), and that as gait distinctiveness increases the distance of person recognition increases as well (Simhi & Yovel, 2019). When viewing a person, we do not know when and how we will see them in the future, and what information may be available for recognition at that time. The fact that all information available contributes to the whole person representation, including the face, body form and, in cases of

dynamic presentation, the gait, aligns with this idea that all of the information about a person may be important for recognition.

Future studies may be able to use the same approach to examine other perceptual judgments such as trait inferences for example, and the extent to which they can be predicted from different components of the whole person – as has already been done for body shape features (Hill et al., 2016). It is possible that for different judgments, particular components (face, body form and gait) might be more important than others. The methods presented here can now be used to test this for other judgements such as gender, age, attractiveness or emotion. In addition, previous works in face recognition (i.e. Butcher & Lander, 2016) and also recognition based on gait (Simhi & Yovel, 2019) have shown that distinctiveness can contribute to our ability to recognize people based on their face or gait, consistent with the relationship that was reported between distinctiveness and memorability (see for example Bainbridge, Isola, & Oliva, 2013). Using data collected in the current study, future studies can examine whether the distinctiveness of particular components of the whole person influences memorability more than others relative to the distinctiveness of the whole person.

Given that people are perceived as a whole in daily life, this study highlights that their representation is not limited just to the face, but is influenced by all the information that is available about them. In examining the basis of person recognition and representation therefore, we must present stimuli that are as close as possible to this naturalistic perception, so that we can truly understand how people are perceived and recognized in real life.

Acknowledgments

We would like to thank Alice O'Toole for providing us with the Video Database of Moving Faces and People. Special thanks to undergraduate students Merav Bilu, Talya Rapp and Guy Brilliant, and to our research assistant Itai Linzen for their assistance in preparing this experiment and collecting the data. The study was funded by an Israeli Science Foundation grant (446/2016) to GY.

Open Practices Statement

The data for all experiments are available at:

https://osf.io/8te3q/?view_only=fbf8a5876ff54c02ab52d871279a2203

The experiments were not preregistered.

References

- Abudarham, N., & Yovel, G. (2016). Reverse engineering the face space: Discovering the critical features for face identification. *Journal of Vision, 16*(3), 40. <https://doi.org/10.1167/16.3.40>
- Aviezer, H., Trope, Y., & Todorov, A. (2012a). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science (New York, N.Y.), 338*(6111), 1225–1229. <https://doi.org/10.1126/science.1224313>
- Aviezer, H., Trope, Y., & Todorov, A. (2012b). Holistic person processing: faces with bodies tell the whole story. *Journal of Personality and Social Psychology, 103*(1), 20–37. <https://doi.org/10.1037/a0027411>

- Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. <https://doi.org/10.1037/a0033872>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face Recognition in Poor-Quality Video: Evidence From Security Surveillance. *Psychological Science*, *10*(3), 243–248. <https://doi.org/10.1111/1467-9280.00144>
- Butcher, N., & Lander, K. (2016). Exploring the motion advantage: evaluating the contribution of familiarity and differences in facial motion. *The Quarterly Journal of Experimental Psychology*, *0218*(February), 1–30. <https://doi.org/10.1080/17470218.2016.1138974>
- Diedenhofen, B., & Musch, J. (2015). cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE*, *10*(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Hahn, C. A., O’Toole, A. J., & Phillips, P. J. (2015). Dissecting the time course of person recognition in natural viewing environments. *British Journal of Psychology (London, England : 1953)*, 1–18. <https://doi.org/10.1111/bjop.12125>
- Hill, M. Q., Streuber, S., Hahn, C. A., Black, M. J., & O’Toole, A. J. (2016). Creating Body Shapes From Verbal Descriptions by Linking Similarity Spaces. *Psychological Science*, *27*(11), 1486–1497. <https://doi.org/10.1177/0956797616663878>
- Hu, Y., Parde, C. J., Hill, M. Q., Mahmood, N., & O’Toole, A. J. (2018). First Impressions of Personality

Traits From Body Shapes. *Psychological Science*, 29(12), 1969–1983.

<https://doi.org/10.1177/0956797618799300>

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.

Larsen, P. K., Simonsen, E. B., & Lynnerup, N. (2008). Gait analysis in forensic medicine. *Journal of Forensic Sciences*, 53(5), 1149–1153. <https://doi.org/10.1111/j.1556-4029.2008.00807.x>

Lee, K., Byatt, G., & Rhodes, G. (2000). Caricature Effects, Distinctiveness, and Identification: Testing the Face-Space Framework. *Psychological Science*, 11(5), 379–385. <https://doi.org/10.1111/1467-9280.00274>

Leopold, D. A., Bondar, I. V., & Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature*, 442(7102), 572–575. <https://doi.org/10.1038/nature04951>

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, 102(45), 16518–16523. <https://doi.org/10.1073/pnas.0507650102>

O'Toole, A. J., Harms, J., Snow, S. L., Hurst, D. R., Pappas, M. R., Ayyad, J. H., & Abdi, H. (2005). A video database of moving faces and people. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5), 812–816. <https://doi.org/10.1109/TPAMI.2005.90>

O'Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: dissecting identity with a fusion approach. *Vision Research*, 51(1), 74–83. <https://doi.org/10.1016/j.visres.2010.09.035>

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences*, 6(6), 261–266. <https://doi.org/10.1016/S1364->

6613(02)01908-3

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. <https://doi.org/10.1163/156856897X00366>

Pilz, K. S., & Thornton, I. M. (2017). Idiosyncratic body motion influences person recognition. *Visual Cognition*, *25*(4–6), 539–549. <https://doi.org/10.1080/13506285.2016.1232327>

Pilz, K. S., Vuong, Q. C., Bühlhoff, H. H., & Thornton, I. M. (2011). Walk this way: approaching bodies can influence the processing of faces. *Cognition*, *118*(1), 17–31. <https://doi.org/10.1016/j.cognition.2010.09.004>

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, *46*(18), 2977–2987. <https://doi.org/10.1016/j.visres.2006.03.002>

Rhodes, G., Jeffery, L., Boeing, A., & Calder, A. J. (2013). Visual coding of human bodies: perceptual aftereffects reveal norm-based, opponent coding of body identity. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(2), 313–317. <https://doi.org/10.1037/a0031568>

Rice, A., Phillips, P. J., Natu, V., An, X., & O’Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological Science*, *24*(11), 2235–2243. <https://doi.org/10.1177/0956797613492986>

Simhi, N., & Yovel, G. (2016). The contribution of the body and motion to whole person recognition. *Vision Research*, *122*, 12–20. <https://doi.org/10.1016/j.visres.2016.02.003>

Simhi, N., & Yovel, G. (2017). The role of familiarization in dynamic person recognition. *Visual Cognition*, *25*(4–6), 550–562. <https://doi.org/10.1080/13506285.2017.1307298>

- Simhi, N., & Yovel, G. (2019). Dissociating identity from gait : A virtual reality study of the role of dynamic identity signatures in person recognition, 1–37.
<https://doi.org/https://doi.org/10.31234/osf.io/39g5k>
- Streuber, S., Quiros-Ramirez, M. A., Hill, M. Q., Hahn, C. A., Zuffi, S., O’Toole, A., & Black, M. J. (2016). Body talk. *ACM Transactions on Graphics*, 35(4), 1–14. <https://doi.org/10.1145/2897824.2925981>
- Valentine, T. (1991). A Unified Account of the Effects of Distinctiveness, Inversion, and Race in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, 43(2), 161–204.
<https://doi.org/10.1080/14640749108400966>
- Valentine, T. (2001). Face-Space Model of Face Recognition. In M. J. Wenger & J. T. Townsend (Ed.), *Computational, geometric, and process perspectives of facial cognition: Contexts and challenges* (pp. 83–113). Mahwah, NJ: Lawrence Erlbaum Associates.
- Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face-Space: A Unifying Concept in Face Recognition Research. *Quarterly Journal of Experimental Psychology*, 69(10), 1996–2019.
<https://doi.org/10.1080/17470218.2014.990392>
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion (Washington, D.C.)*, 7(3), 487–494. <https://doi.org/10.1037/1528-3542.7.3.487>
- Yovel, G., & O’Toole, A. J. (2016). Recognizing People in Motion. *Trends in Cognitive Sciences*, 20(5), 383–395. <https://doi.org/10.1016/j.tics.2016.02.005>
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, 12(4), 399–413. <https://doi.org/10.1037/1082-989X.12.4.399>