Journal of Experimental Psychology: Learning, Memory, and Cognition

Learning Faces as Concepts Rather Than Percepts Improves Face Recognition

Linoy Schwartz and Galit Yovel

Online First Publication, December 20, 2018. http://dx.doi.org/10.1037/xlm0000673

American Psychological Association

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/18/$12.00

http://dx.doi.org/10.1037/xlm0000673

Learning Faces as Concepts Rather Than Percepts Improves Face Recognition

Linoy Schwartz and Galit Yovel
Tel Aviv University

Our ability to recognize familiar faces is remarkable. During the process of becoming familiar with new people we acquire both perceptual and conceptual information about them. Which of these two types of information contributes to our ability to recognize a person in future encounters? Previously, we showed that associating faces with person-related conceptual information (e.g., name, occupation) during learning improves face recognition. Here, we provide further evidence and assess several possible accounts to the *conceptual encoding benefit* in face recognition. In a series of experiments, participants were asked to make perceptual (e.g., how round/symmetric is the face?) or conceptual (e.g., how trustworthy/intelligent does the face look?) evaluations about faces. We found better face recognition following conceptual than perceptual encoding. We further showed that this effect cannot be attributed to more global than part-based feature processing, more variable ratings, or more elaborative encoding during conceptual than perceptual evaluations. Finally, we showed that the conceptual over perceptual encoding advantage reflects a conceptual encoding benefit rather than a perceptual encoding cost. Overall these findings show that conceptual evaluations do not improve recognition by modifying the perceptual representation of a face (e.g., elaboration, global processing). Instead, we propose that face recognition benefits from representing faces as socially meaningful concepts rather than percepts during learning. These results highlight the importance of linking cognition and perception to understand recognition.

Keywords: concepts, face recognition, inferred traits, perceptual learning, social cognition

Supplemental materials: http://dx.doi.org/10.1037/xlm0000673.supp

Face recognition involves matching an incoming face to a representation of a previously seen face that we store in memory. What is the nature of the representation that we generate for familiar faces in memory? A prominent model of face recognition (Bruce & Young, 1986) suggests that familiar faces are represented in a view-invariant, abstract manner in a *face recognition unit*. This face recognition unit is separated from other sources of information about the identity of a familiar person such as their voice, name, or any other semantic information about them. This model therefore suggests that the visual representation of a familiar face in memory is separated from the conceptual information associated with it. Accordingly, many studies that examined the representation of familiar faces have focused on their perceptual representation, overlooking the role of conceptual information during the process of becoming familiar with a new face. In particular, these studies suggested that the extensive perceptual exposure that we have with different images of familiar faces leads to the generation of an invariant perceptual representation, in which image-based information is disregarded, and perceptual information that is inherent to the identity of the face remains (Burton, 2013; Burton, Jenkins, Hancock, & White, 2005; Burton, Jenkins, & Schweinberger, 2011). These studies, however, did not examine the possible contribution of conceptual information during face learning to face recognition.

In addition to the rich perceptual information that we acquire about a familiar person during the process of familiarization, faces are associated with semantic, emotional, and/or episodic information. To study the contribution of conceptual information to face recognition, Schwartz and Yovel (2016) presented subjects with a face learning task in which participants associated a face with a person-related label (i.e., person name, occupation) to generate a conceptual representation. Results showed that associating faces with person-related labels improved face recognition relative to faces that were not associated with any label. Importantly, the effect of labeling was not found when faces were associated with person-unrelated labels (e.g., object names, symbols), indicating that the effect of associating person-related labels to faces was not attributable to individuation of or attention to the labeled faces. Instead, it is the generation of a conceptual representation during learning that improves later recognition. These findings are consistent with earlier studies that showed that associating faces with

Linoy Schwartz, Sagol School of Neuroscience, Tel Aviv University; Galit Yovel, School of Psychological Sciences and Sagol School of Neuroscience, Tel Aviv University.

Data from this article were presented in scientific meetings and university colloquia. Data files of the five experiments can be found in https://osf.io/dnye2/.

Correspondence concerning this article should be addressed to Galit Yovel, School of Psychological Sciences and Sagol School of Neuroscience, Tel Aviv University, or to Linoy Schwartz, Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: gality@post.tau.ac.il or linoyshw@mail.tau.ac.il

names improves recognition of faces from other races or other ages (Gordon & Tanaka, 2011; Yovel et al., 2012). Taken together, these findings suggest that associating faces with person-related information improves face recognition.

The current study aims to further explore the conceptual benefit in face recognition. Whereas in our previous studies we examined the effect of associating person-related (e.g., name) and person-unrelated (e.g., object) information with faces, here we aimed to study the conceptual benefit in face recognition by comparing it with other types of information that can be extracted from the face, but are not expected to generate a conceptual representation. When we see a face we can describe it based on its perceptual features, such as how large are the eyes or how thick are the lips. In addition, humans make trait inferences (e.g., trustworthy, dominant) about faces based on their facial features (Oosterhof & Todorov, 2008). However, despite the fact that trait inferences and perceptual evaluations about faces are both based on the same physical information, they may contribute differently to face recognition. By asking subjects to make explicit trait evaluations, we encourage the transformation of the image of a face from a perceptual to social-conceptual representation. Such transformation is not expected to occur when making explicit perceptual evaluations.

The effect of conceptual and perceptual evaluations in face recognition has been studied more than two decades ago (Bower & Karlin, 1974; Courtois & Mueller, 1979; Mueller, Carlomusto, & Goldstein, 1978; Strnad & Mueller, 1977; Winograd, 1976, 1981) to examine the role of level of processing (LOP) on episodic memory (Craik & Lockhart, 1972; Craik & Tulving, 1975). The LOP effect was originally shown for words. Participants were asked to learn a list of words, while making semantic evaluations (a.k.a. deep encoding) or physical evaluations such as the color of the fonts (a.k.a. shallow encoding) about words. Whereas the question of what constitutes deep or shallow encoding remained controversial (Baddeley, 1978; Eysenck, 1978; Nelson, 1977), results showed that recognition was better for words that underwent semantic than physical evaluations. To extend these findings to faces, other studies asked participants to make trait inferences (e.g., how intelligent does the face look?) or perceptual evaluations about visual features of a face (e.g., how large are the eyes?). It is noteworthy that unlike words in which the semantic meaning is independent from the physical information of the stimulus, both trait inferences and perceptual evaluations rely on the same physical facial information. Still, results revealed better recognition for face images that underwent trait than perceptual evaluations, and were taken to support the level of processing approach (Craik & Lockhart, 1972; Craik & Tulving, 1975).

Whereas these findings may seem consistent with a conceptual benefit in face recognition, other mechanisms have been suggested to account for this effect. In particular, rather than generating a conceptual representation, trait inferences may modify the perceptual representation of faces in one of the following ways: First, conceptual processing may involve more elaborative processing of facial information. For example, Winograd (1981) has suggested that trait inferences may involve the extraction of more facial features than perceptual evaluations (*The feature elaboration hypothesis*). This account predicts that conceptual evaluations may take longer than perceptual evaluations during encoding, as more features are being encoded. Second, previous studies that compared trait and perceptual evaluations required subjects to make part-based judgments about specific facial features during the perceptual evaluation condition. However, conceptual processing may involve more global than part-based face processing (Coin & Tiberghien, 1997). This account predicts that making perceptual evaluations about global facial features may abolish the conceptual over perceptual benefit in face recognition. Third, conceptual evaluations may generate more variable ratings across faces relative to more narrow perceptual evaluations. More variable ratings may increase the perceptual distance between the representations of faces (i.e., individuation) during face learning following conceptual relative to perceptual evaluations. Finally, the conceptual over perceptual benefit may in fact reflect a perceptual cost rather than a conceptual benefit. That is, perceptual evaluations may interfere with face recognition relative to faces encoded with no evaluations (e.g., Schooler & Engstler-Schooler, 1990), whereas conceptual evaluations may not improve face recognition relative to faces that are encoded with no evaluations. All these different explanations are inconsistent with the idea that face recognition benefits from the generation of a conceptual representation for faces during learning, in which faces are no longer represented as pure percepts but as concepts. Instead, these accounts suggest that conceptual evaluations modify the perceptual representation of faces in a way that benefits recognition.

Another limitation of previous studies that compared conceptual and perceptual evaluations was that the same face images that were presented during learning were also presented during the recognition phase. Thus, these studies did not examine recognition of face identity but recognition of the image of the studied face. A conceptual account of face recognition predicts that the benefit of conceptual encoding would not be limited to the image presented at study, but would improve recognition of the identity of a face and generalize to different images of the same individual. Our previous findings indeed show that associating person-related information with faces improves recognition of different images of the same identity (Schwartz & Yovel, 2016). It is therefore important to establish whether the benefit of making conceptual than perceptual evaluations for face recognition generalizes to different images of the learned identities.

To address all these different accounts for the *conceptual encoding benefit* in face recognition, we conducted a series of experiments that systematically examined the effect of conceptual and perceptual evaluations during learning on face recognition. In all the experiments, participants underwent a learning phase during which they were asked to make trait inferences or perceptual evaluations about each face. During the following recognition phase, different images of the learned identities and new identities were presented, and participants were asked to indicate whether they recognized each face from the learning phase or whether that was a new face (see Figure 1).

In Experiment 1 we asked whether conceptual evaluations (i.e., trait inferences) would enhance face recognition more than perceptual evaluations for new images of the same identity. We also examined reaction time (RT) during the learning phase, to reveal whether conceptual evaluations require more processing time, as the feature elaboration hypothesis predicts (see also Experiment 5). In Experiment 2, we examined whether the conceptual over perceptual benefit is due to more global processing of facial features. In Experiment 3, we examined whether the variance of conceptual rating is larger than

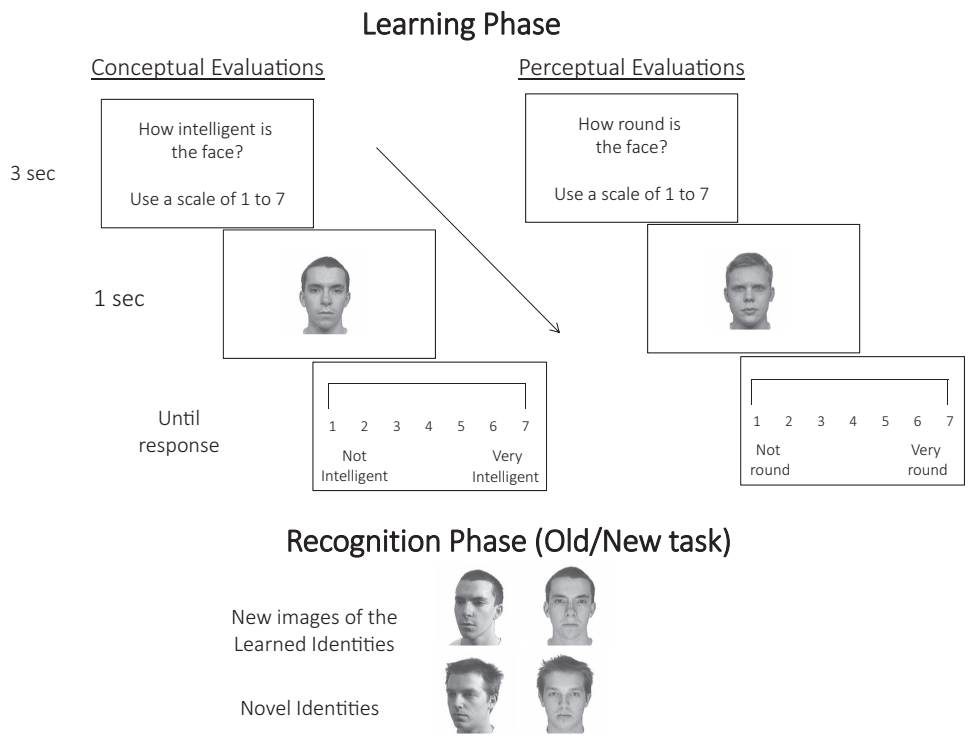

Figure 1. Experimental Design of Experiments 1 and 2: During the learning phase, participants rated 20 different identities based on either their perceptual appearance or their inferred personality traits. A total of five questions, either perceptual or conceptual, were presented about each face. Faces appeared in miniblocks by identity with each face repeating five times, each time following a different question (the figure shows only one question for each condition out of the set of five questions). During the recognition phase, different images of the learned faces and novel faces were presented. The images of the learned identities differed either in lighting or in lighting and view from the learning phase faces, and participants were asked to make an old/new decision. These images are used with permission by the Vision Science Lab.

the variance of perceptual rating, which may lead to better discrimination among the learned faces. In Experiments 4 and 5, we assessed whether better performance following conceptual than perceptual evaluations, reflects a conceptual evaluation benefit or a perceptual evaluation cost, by comparing performance to faces presented with no evaluations. Table 1 summarizes the four different hypotheses that were tested in the five experiments.

Experiment 1

The first goal of Experiment 1 was to assess whether the conceptual over perceptual benefit is generalized to new images of the same identities presented during the learning phase. To that effect, we presented at study frontal faces and at test images of these identities that differ either in view or in view and lighting. These test images were presented together with new identities and participants were asked to make an old-new decision. In addition, we examined whether RT to make perceptual evaluations is faster than for making conceptual evaluations, as the feature elaboration hypothesis predicts.

Participants were asked to make five trait inferences or five perceptual evaluations for each identity. To maximize the benefit of perceptual evaluations, we chose a subset of features that were shown to be important for the identity of the face (Abudarham & Yovel, 2016). For trait inferences we selected traits that have been commonly used in previous studies that examined the association between trait inferences and facial features (e.g., Oosterhof & Todorov, 2008). Accuracy and RTs were measured during the recognition phase, and RT was measured for the conceptual and perceptual evaluations during the learning phase.

Table 1
Four Alternative Explanations to the Conceptual Benefit Hypothesis Were Tested in a Series of Five Experiments

Hypothesis	Experiment
Feature elaboration hypothesis	Experiments 1–5
Global processing hypothesis	Experiment 2
Rating variability hypothesis	Experiment 3
Perceptual cost hypothesis	Experiments 4–5

Method

Participants. Twenty-nine undergraduate students from Tel Aviv University (Caucasians, 17 females, between the ages 19–28, mean age = 23.06, *SD* = 2.56) participated in this experiment. All participants received course credit or payment for participating in the experiment and reported normal or corrected to-normal vision.

The study was approved by the ethics committee of Tel Aviv University, and all participants gave informed consent to participate in the study. Because previous studies that used a similar paradigm did not report measures of variance, we could not perform power analysis based on their data. We therefore used a sample size that is similar to our previous studies using an old-new task with the same stimuli (Schwartz & Yovel, 2016), and based on data we obtained in Experiment 1 we performed power analysis to determine minimal sample size needed for the following experiments.

Stimuli. Stimuli were grayscale face photographs of 40 Caucasian young adult males, all taken from the Harvard face database (Duchaine & Nakayama, 2006). The race of the stimuli matched the race of the participants. All faces were adult males with no glasses, facial hair, long hair, or scars. All photos were centered and their backgrounds were removed. From the pool of 40 identities, 20 identities were randomly assigned at the beginning of each experiment to be presented in the learning phase. The remaining 20 identities were novel faces: faces that would appear only during the recognition phase of the experiment as novel faces.

Of the 20 identities that were randomly chosen for each participant to be learned during the learning phase, 10 identities were randomly assigned to the perceptual condition and the other 10 identities were assigned to the conceptual condition.

During the learning phase frontal view faces were presented. During the recognition phase, faces that differed in either lighting or in both lighting and view were presented (see Figure 1).

Procedure. The experiment included a 20-min leaning phase during which participants were asked to learn the identities of faces for a later recognition test, and while studying the faces to make five conceptual or perceptual judgments about each of the 20 learned faces. The learning phase was immediately followed by a 5-min recognition phase.

Learning phase. During the learning phase, participants were asked to learn a total of 20 different face identities and rate them based on either their perceptual appearance or their inferred personality traits. Each trial began with either a perceptual or a conceptual question from a set of five questions. Conceptual and perceptual miniblocks of five questions were presented in a random order. Each question appeared on the screen for 3 second, followed by a 1-second interstimulus interval after which a face was presented for 1 second. Following the face offset, a scale of 1 to 7 appeared on the screen, and the participants were asked to rate the face based on the question on that scale (see Figure 1).

The total of five perceptual or conceptual questions were presented sequentially one after the other for each identity. Thus, each face was presented five times, following each one of the five different questions. The order of the five questions on each miniblock was random. The five conceptual questions were "How intelligent/dominant/trustworthy/friendly/aggressive is the following face?" The five perceptual questions focused on specific facial features that were found to play a significant role in face recognition (Abudarham & Yovel, 2016): "How round are the eyes?"; "How dark are the eyes?"; "How thick are the lips?"; "How thick are the eyebrows?"; "How long is the hair?"

Recognition phase. The recognition phase included an old–new task. Participants were presented with a series of 80 images of 40 identities. Half of the identities were learned during the learning phase (evaluated based on either conceptual or perceptual features), and the other half were novel identities. The test images were presented sequentially, each one for 1 second in a random order. Two different images were used for each facial identity throughout the recognition phase: one image differed in lighting and the other in both lighting and view from the learned images, and similar images were used for the novel faces (see Figure 1). Before the recognition phase commenced, participants were told that new images of the learned identities would be presented during the following phase, and that their goal was to recognize the identity of the face rather than the specific image that was presented during the learning phase. Participants were asked to press one key for old faces, regardless of whether the faces were learned with conceptual or perceptual questions, and a different key for new faces. After the participant made a response, the next trial commenced.

Data analysis.

Recognition phase. For all the experiments reported, we used two measures to estimate performance level on the task: The mean sensitivity (d') for discriminating between old and novel faces and the percent of correct responses for old and new faces. Hit rate of 1 was converted to $1 - (1/(2n)$; Hit rate of 0 was converted to $1/(2n)$ where n is the number of trials. Both measures revealed similar findings. We report statistical tests on the d' measures, and report in tables the average Hit rates and False Alarm rates and in supplementary tables d' and RT measures for each condition. The sensitivity measure, d' was calculated based on the same false alarm rate for the perceptual and conceptual conditions. Therefore, d' measures correspond to the Hit rate but are reported as d' to provide a measure that takes into account also the false alarm rate on this task.

Reaction times (RTs) were examined during the recognition phase from the onset of the face until participants made an old/new decision and were calculated only for correct trials. To remove outliers, trials with RTs that were longer than 10 seconds or shorter than 200 msec. were excluded from analysis. Following this exclusion, trials with RTs that were 2.5 standard deviations from the mean of each participant were removed from analysis. This trimming procedure resulted in exclusion of an average of 3.3% of the correct trials.

Learning phase. RTs for making perceptual and conceptual evaluations were measured for each of the five conceptual and five perceptual evaluations from the onset of a face until a participant rated the face. The average RT across the five evaluations of each condition was computed for each participant. Similar to the recognition phase, RTs longer than 10 seconds and shorter than 200 ms were removed, followed by exclusion of RTs that were smaller/larger than 2.5 standard deviations from the mean of each participant. Because there were no correct or incorrect responses on this part of the experiment, all trials were included in this analysis.

Results

Recognition phase.

Accuracy. A repeated measures ANOVA on sensitivity scores (d') with Type of Evaluation (Conceptual, Perceptual) and Test face view (different lighting, different lighting & view) as within-subject factors, revealed a significant advantage for recognizing conceptually over perceptually learned faces, $F(1, 28) = 43.71$, $p < .001$, $\eta_p^2 = 0.61$. Recognition was better for faces that differed

from the learned images only in lighting as compared to faces that differed in both lighting and view, $F(1, 28) = 13.22$, $p = .001$, $\eta_p^2 = 0.32$. No interaction between Type of Evaluation and Test face view was found ($F(1, 28) < 1$; Figure 2A; see Table 2 for Hits and False alarm rates, and Supplementary Table 1 for d' and RTs).

Reaction time. A repeated measures ANOVA on RTs during the recognition phase with Type of Evaluation (Conceptual, Perceptual) and Test face view (different lighting, different lighting and view) as within-subject factors revealed marginally significant shorter RTs for recognizing conceptually than perceptually evaluated faces $F(1, 28) = 3.96$, $p = .056$, $\eta_p^2 = 0.12$, but no significant difference in RTs for recognizing faces that differed in lighting as compared to faces that differed in both lighting & view, $F(1, 28) < 1$. The interaction between Type of Evaluation and Test face view was not statistically significant, $F(1, 28) = 2.65$, $p = .12$, $\eta_p^2 = 0.09$.

Learning phase. A paired-sample t test on RTs for making conceptual and perceptual evaluations revealed significantly shorter RTs for conceptual than perceptual evaluations, $t(28) = 6.2$, $p < .0001$, Cohen's $d = 1.18$. Table 3 reports the mean and standard deviation of this measure for all five experiments.

Discussion

Our results show that making conceptual evaluations about faces during learning improves recognition of new images of the learned faces more than perceptual evaluations. These findings go beyond previous reports by showing that the benefit of making trait inferences than perceptual evaluations about faces is not limited to the face images presented during the learning phase but is generalized to new images of the learned identities that differ in view and lighting.

Furthermore, the benefit of conceptual evaluations cannot be attributed to longer processing time during the learning phase, as RTs were shorter for conceptual than perceptual evaluations. These results are inconsistent with the idea that the *conceptual encoding benefit* is attributable to more elaborated processing of faces during encoding (Winograd, 1981). Instead, it shows that making conceptual evaluations is a more efficient type of face processing then judging facial features and is consistent with the idea that conceptual encoding is important for face recognition (Schwartz & Yovel, 2016; Yovel et al., 2012).

It is possible, however, that the benefit of the conceptual over perceptual evaluations may not be attributable to the number of features that are processed during encoding, but the type of features used during conceptual than perceptual encoding. In our study, as well as in previous studies (e.g., Courtois & Mueller, 1979; Mueller et al., 1978; Winograd, 1976, 1981), perceptual evaluations focused on facial features that may encourage more local-based feature processing in contrast to conceptual evaluations that may be based on more global impressions (Coin & Tiberghien, 1997). We therefore conducted the same experiment changing only the perceptual condition: instead of evaluating local facial features such as the eyes or the lips, participants were asked to evaluate faces based on more global features such as face symmetry, width or skin tone. The experiment was therefore identical to Experiment 1, with the only difference that participants were asked to evaluate global facial features rather than local facial features.

Experiment 2

Method

Participants. Power analysis based on the effect size obtained in Experiment 1 indicated that 10 participants are needed to obtain a significant conceptual benefit in face recognition. To have a more balanced design when comparing among the different experiments, we kept the sample size to $n = 25$–30 across all experiments.

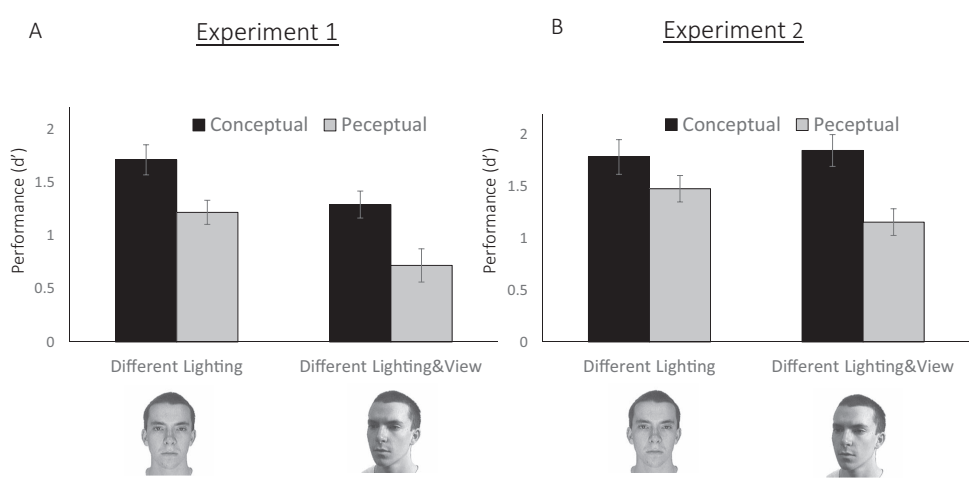

Figure 2. (A) Recognition level (d') in Experiment 1 for faces that differed in lighting or lighting and view from the learned faces, following perceptual or conceptual evaluations. Perceptual evaluations were based on local facial features such as "how thick the eyebrows/lips are?" (B) Recognition level (d') in Experiment 2 for faces that differed in lighting or lighting & view from the learned faces, following perceptual or conceptual evaluations. Perceptual evaluations were based on global facial features such as "how symmetric/wide is the face?". These images are used with permission by the Vision Science Lab.

Table 2
Mean, SD, and 95% Confidence Interval (CI) for Hit Rate and False Alarm Rate for the Conceptual and Perceptual Conditions in Experiment 1

	Hit rate				False alarm rate	
	Conceptual evaluations		Perceptual evaluations			
Measure	Lighting	View and lighting	Lighting	View and lighting	Lighting	View and lighting
Mean	.78	.64	.62	.44	.24	.24
SD	.16	.18	.18	.26	.16	.18
95% CI−	.72	.58	.56	.35	.18	.17
95% CI+	.84	.71	.69	.54	.30	.31

Twenty-five undergraduate and graduate students from Tel Aviv University (21 females, between the ages 19–27, mean age = 22.84, *SD* = 2.44) participated in this experiment. All participants received course credit or payment for participating in the experiment and reported normal or corrected to-normal vision. The study was approved by the ethics committee of Tel Aviv University, and all participants gave informed consent to participate in the study.

Procedure. The procedure was identical to the procedure described in Experiment 1, with the only difference in type of perceptual questions. In this experiment, in order to encourage global processing of the face, the perceptual questions were "how symmetric/round/wide/ was the face or how smooth/bright was the skin tone of the face?"

Results

Recognition phase.

Accuracy. A repeated measure ANOVA on sensitivity scores (d') with Type of Evaluation (Conceptual, Perceptual) and Test face view (different lighting, different lighting & view) as within-subject factors revealed a significant advantage for recognizing conceptually over perceptually learned faces, $F(1, 24) = 36.65$, $p < .001$, $\eta_p^2 = 0.6$. There was no difference in recognizing faces that differed only in lighting as compared to faces that differed in both lighting and view, $F(1, 24) = 1.27$, $p > .25$, $\eta_p^2 = 0.05$, but the interaction between Type of Evaluation and Test face view was significant, $F(1, 24) = 6.22$, $p = .02$, $\eta_p^2 = 0.2$. Post hoc Tukey's tests indicate that the benefit of conceptual over perceptual evaluations was significant both for test faces that differed in lighting and view ($p < .001$, Cohen's $d = 1.1$), and for faces that differed only in lighting from the learned images ($p = .02$, Cohen's $d = 0.67$) but was larger for the former than the latter (Figure 2B).

Table 3
Mean (Standard Deviation) for Reaction Time (ms) for Conceptual and Perceptual Evaluations During the Learning Phase in Each of the Five Experiments

Experiment	Conceptual evaluation	Perceptual evaluation
Experiment 1	2020.1 (709.98)	2300.31 (657.56)
Experiment 2	2319.18 (656.91)	2645.55 (618.9)
Experiment 3	2439.14 (596.22)	2608.75 (582.01)
Experiment 4	2428.32 (445.97)	2594.14 (511.46)
Experiment 5	2280.22 (704.08)	2759.74 (897.15)

Reaction time. A repeated measure ANOVA on RTs with Type of Evaluation (Conceptual, Perceptual) and Test face view (different lighting, different lighting and view) as within-subject factors revealed no significant difference in RT for recognizing conceptually than perceptually learned faces, $F(1, 24) = 2.44$, $p = .13$, $\eta_p^2 = 0.09$. Reaction time was faster for recognition of test faces that differed only in lighting as compared to faces that differed in both lighting and view, $F(1, 24) = 5.00$, $p = .035$, $\eta_p^2 = 0.17$. A significant interaction between Type of Evaluation and Test face view was also found, $F(1, 24) = 5.82$, $p = .024$, $\eta_p^2 = 0.2$. Post hoc Tukey's test revealed that this interaction indicates faster RT for conceptually than perceptually judged faces when test faces differed only in lighting ($p = .027$, Cohen's $d = 0.68$), but not when test faces differed in both lighting & view from the learned faces ($p > .25$, Cohen's $d = 0.06$). (see Table 4 for Hit and False alarm rates, and Supplementary Table 2 for d' and RTs).

Learning phase. A paired-sample t test on RTs for making conceptual and perceptual evaluations revealed significantly shorter RTs for conceptual than perceptual evaluations, $t(24) = 6.4$, $p < .0001$, Cohen's $d = 1.25$ (see Table 3).

Comparing performance for local and global perceptual evaluations. Experiments 1 and 2 used different types of perceptual evaluations. To examine whether the type of perceptual evaluations (local versus global) influenced the level of face recognition, we compared recognition performance for faces that underwent perceptual evaluations in Experiment 1 and 2 and examined whether there was any difference in recognition following global versus local-based evaluations. A mixed ANOVA on sensitivity scores (d') with Test face view (different lighting, different lighting and view) as a within-subject factor, and Experiment (Experiment 1 – local perceptual evaluations, Experiment 2 – global perceptual evaluations) as a between-subjects factor revealed a marginally significant effect for Experiment, $F(1, 52) = 4.00$, $p = .05$, $\eta_p^2 = 0.07$, reflecting numerically higher performance for global than local evaluations, and no interaction between Test face view and Experiment, $F(1, 52) < 1$. Thus, although global perceptual evaluations yielded slightly higher recognition rates, performance was still much better for conceptual evaluations. In the following experiments, we kept using the global perceptual evaluations.

Discussion

Experiment 2 revealed a significant advantage for recognizing faces following conceptual than perceptual evaluations also when

Table 4
Mean, SD, and 95% Confidence Interval (CI) for Hit and False Alarm Rates for the Conceptual and Perceptual Conditions in Experiment 2

	Hit rate				False alarm rate	
	Conceptual evaluations		Perceptual evaluations			
Measure	Lighting	View and lighting	Lighting	View and lighting	Lighting	View and lighting
Mean	.76	.78	.66	.54	.18	.20
SD	.18	.17	.17	.20	.13	.18
95% CI−	.70	.71	.59	.46	.13	.13
95% CI+	.83	.84	.72	.62	.23	.27

the perceptual evaluations relied on global facial features. These findings add to Experiment 1 and to previous reports that involved perceptual evaluations of local facial features (Bower & Karlin, 1974; Courtois & Mueller, 1979; Mueller et al., 1978; Strnad & Mueller, 1977; Winograd, 1976, 1981) by showing that better recognition following conceptual than perceptual evaluations is not due to global versus local face processing, respectively. Overall these findings provide further support for the benefit of making conceptual evaluations during face learning on face recognition of new images of the learned identities. In Experiment 2, we also replicated our finding of faster RT for making conceptual and perceptual evaluations, indicating that better recognition was not due to more elaborated processing of faces during conceptual judgments (see Table 3).

Another difference between conceptual and perceptual encoding that may account for the *conceptual over perceptual encoding benefit* in face recognition is that trait evaluations may yield more variable responses than evaluations of perceptual features. For example, participants may use more variable ratings between faces for indicating whether faces look friendly (e.g., some faces may look very friendly and others less friendly) or intelligent than whether faces look symmetric (e.g., all faces may look similarity symmetric) or round. Such larger variability may increase the perceptual distance between the learned representations following conceptual than perceptual representation and contribute to improved face recognition. To address this possible explanation, in Experiment 3 we examined the variability of conceptual and perceptual evaluations.

Another alternative explanation to the *conceptual benefit* in face recognition is greater redundancy in the type of evaluations that were made in the perceptual than the conceptual condition. Greater redundancy among the different questions used for a given condition may result in less elaborated encoding of facial information and may lead to lower recognition. To examine this possibility, we compared the correlations across conceptual and perceptual evaluations. This allowed us to assess whether better performance for conceptually than perceptually evaluated faces may be due to greater overlap in the type of questions used in perceptual than conceptual evaluations.

Experiment 3

The main goals of Experiment 3 were to compare the variability of conceptual and perceptual rating across faces, and to assess the correlations across the different judgments for each type of evaluations. In Experiments 1 and 2, we randomly allocated the 40 different face identities to the learning or recognition phase and within the learning phase to the conceptual or perceptual conditions across participants. Thus, a relatively small and variable number of participants rated each of the faces conceptually or perceptually (n = 2–11). This does not allow reliable comparison between the two types of evaluations across faces.

To obtain a more stable and reliable measure of the variability of conceptual and perceptual ratings across faces we performed a separate rating experiment. Participants were asked to make the same conceptual and perceptual evaluations that they were asked to make in Experiment 2, but for all the faces included in the experiment. This allowed us to directly compare the variance of conceptual and perceptual evaluations across faces, for the same set of faces with the same number of participants for each face and type of evaluation as well as to assess the degree of overlap between the conceptual questions relative to perceptual questions by computing the correlations among conceptual or perceptual evaluations.

Method

Participants. Twenty-two participants (between the ages 19–42, mean age = 29.21, *SD* = 7.87) participated in this experiment. The experiment was conducted online on an Israeli survey platform named "Panel4all," similar to Amazon's "Mechanical Turk." All participants received payment for participating in the experiment ($4 for a 30-min experiment). The study was approved by the ethics committee of Tel Aviv University, and all participants gave informed consent to participate in the study.

Procedure. The experiment included a 40-min evaluation phase during which participants were asked to make five conceptual or five perceptual judgments about each of the 40 faces that were presented to them. Similar to the learning phase of the recognition experiments, each question appeared on the screen for 3 s, followed by a 1-s interstimulus interval after which a face was presented for 1 s. Following the face offset, a scale of 1 to 7 appeared on the screen, and the participants were asked to rate the face based on the question on that scale. There were two versions of the experiment: in the first version half of the faces were accompanied with perceptual questions, and the other half with conceptual questions, and vice versa in the second version. Participants were randomly allocated to each of these two versions.

Table 5
Mean and Standard Deviation Across the 40 Faces for Each of the Conceptual and Perceptual Evaluations

	Perceptual evaluation				
Measure	Symmetric	Smooth	Wide	Skin color	Round
Mean	3.93	3.82	4.03	4.13	3.69
SD	.70	.85	1.00	.75	1.05
	Conceptual evaluation				
	Intelligent	Friendly	Aggressive	Dominant	Trustworthy
Mean	3.70	3.52	3.75	3.83	3.69
SD	.62	.86	.86	.58	.68

Results

Table 5 shows the mean and standard deviation for each of the conceptual and perceptual evaluations across faces (see also Figure 3). Examination of the numerical values show no evidence that the conceptual evaluations are more variable than the perceptual evaluations. In fact, the largest variance was for two perceptual evaluations (wide and round) whereas the variability of the conceptual evaluations was overall lower (see Table 5). To examine equality of variance among the 10 different evaluations we performed an F test among all pairs of the 10 different evaluations and corrected for 45 comparisons ($p = .001$). The variance for dominance was significantly lower than the variance for wide and round evaluations ($p < .001$). None of the other evaluations were statistically different from other evaluations. As Table 5 shows, the standard deviation of ratings across faces for the conceptual questions was not numerically larger than the perceptual questions. In fact, most perceptual evaluations were more variable than conceptual evaluations.

To test whether perceptual evaluations were more redundant than conceptual evaluations, we measured the correlations across judgments for both conceptual and perceptual evaluations (see supplementary Table 3). We calculated the average correlation (absolute value) among perceptual and conceptual evaluations. Results show a numerically larger correlation among the conceptual evaluations ($r = .43$) than the perceptual evaluations ($r = .27$), with no significance difference between them, $t(9) = 1.33$, $p = .215$, Cohen's $d = 0.42$. Thus, the conceptual benefit in face recognition is not due to smaller overlap among conceptual than perceptual evaluations.

Finally, we measured RTs for conceptual and perceptual questions. A paired-sample t test on RTs of conceptual and perceptual evaluations, revealed significantly shorter RTs for making conceptual than perceptual evaluations, $t(21) = 3.56$, $p = .002$, Cohen's $d = 0.76$, replicating results from the previous experiments (see Table 3).

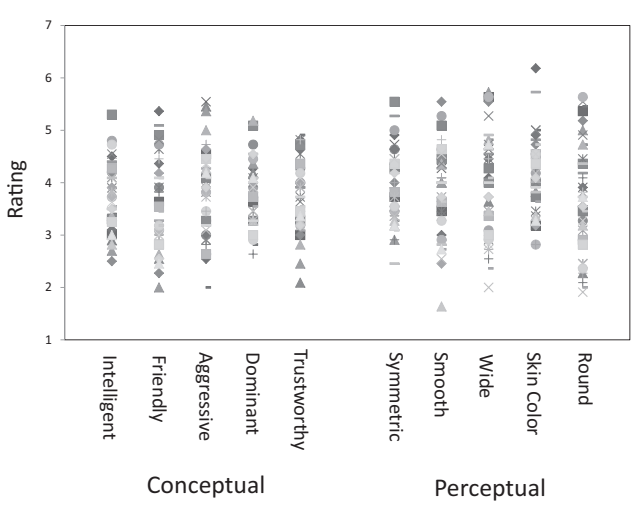

Figure 3. The variance across faces in ratings for the five conceptual and five perceptual questions.

Discussion

Our results show no evidence for larger variability for conceptual than perceptual rating across faces. These results therefore indicate that differences in rating variability cannot account for the conceptual over perceptual benefit in face recognition. Additionally, results indicate that correlations across evaluations were not significantly different between conceptual than perceptual evaluations, thus suggesting that difference in overlap between perceptual and conceptual evaluations also cannot accounts for the recognition findings. Taken together with our previous results, we found no evidence that conceptual evaluations lead to more elaborated processing of facial information than perceptual evaluations.

However, experiments so far that have demonstrated a robust conceptual over perceptual evaluation benefit, have not shown directly that this effect is necessarily attributable to a *conceptual evaluation benefit*. In fact, the conceptual over perceptual benefit may merely reflect a perceptual evaluation cost, that is, making perceptual evaluations may interfere with face encoding, with no evidence for a conceptual benefit. Previous studies have shown that verbal description of facial features (i.e., verbal overshadowing) may interfere with face recognition (Schooler & Engstler-Schooler, 1990). To address this possibility, we conducted the same experiment we ran in Experiment 2 but added a third condition in which participants passively viewed faces while making *no evaluations* about them. If the conceptual over perceptual benefit is merely due to perceptual interference, we expect lower performance when making perceptual evaluations than when making no evaluations and no difference between conceptual evaluations and no evaluations. If instead conceptual evaluations facilitate face recognition, we expect better recognition for faces that underwent conceptual evaluations than for faces that were learned with no evaluations.

Experiment 4

The goal of Experiment 4 was to assess whether the *conceptual over perceptual benefit* in face recognition reflects a conceptual encoding benefit or a perceptual encoding cost. To that effect, we asked participants to either make conceptual evaluations, perceptual evaluations or no evaluations while studying faces. We then compared recognition level for different images of the learned faces (same test images as in Experiments 1 and 2) across these three conditions.

Method

Participants. Thirty-one participants (17 females, between the ages 18–45, mean age = 29.42, *SD* = 8.34) participated in this experiment. The experiment was conducted online on a survey platform named "Panel4all," similar to Amazon's "Mechanical Turk." All participants received payment for participating in the experiment ($4 for a 30-min experiment). The study was approved by the ethics committee of Tel Aviv University, and all participants gave informed consent to participate in the study.

Procedure.

Stimuli. From the pool of 40 identities, 24 identities were randomly assigned at the beginning of each experiment to be presented during the learning phase. The remaining 16 identities were novel faces: faces that appeared only during the recognition phase of the experiment as novel, previously unseen faces.

Learning phase. The learning phase of this experiment was similar to the procedure described in Experiment 2, with several changes that were added to include the No Evaluation condition (passive viewing with no evaluation). First, in this experiment only eight faces were learned in each condition (Conceptual, Perceptual, No Evaluations), instead of the 10 faces per condition that were learned in previous experiments. Thus, a total of 24 identities were presented during the learning phase, as compared with the 20 identities in the previous experiments.

The learning procedure of the conceptual and perceptual faces was identical to the procedure described in Experiment 2 and included perceptual questions about global facial aspects. For the no evaluation condition, each trial began with a 3-second pause in which participants viewed a blank screen that was followed by the appearance of the face for 1 second. After the face offset the screen was left blank for an additional 1-second period. These durations were chosen to match the encoding duration of the perceptually and conceptually learned faces. Similar to Experiments 1 and 2, a total of five sequential trials were used to study each identity.

Recognition phase. The procedure of the recognition phase was similar to that used in Experiments 1 and 2. A total of 80 faces were presented during the recognition phase. From the 80 face images, in this Experiment 48 were images of the learned faces, presented in different lighting or in different lighting and view from the learned images, and 32 images were two images of 16 novel identities.

Results

Recognition phase.

Accuracy. A repeated measure ANOVA on the sensitivity score (d') with Type of Evaluation (Conceptual, Perceptual, No Evaluations) and Test face view (different lighting, different lighting and view) as within-subject factors revealed a main effect for Type of Evaluation ($F(2, 60) = 9.23$, $p < .001$, $\eta_p^2 = 0.24$). Post hoc Tukey's tests revealed better recognition of the conceptually evaluated faces than the perceptually evaluated faces ($p = .007$, Cohen's $d = 0.66$), and faces presented with no evaluations ($p < .001$, Cohen's $d = 0.67$). No difference was found between faces that were presented with perceptual evaluations and no evaluations ($p > .25$). Also, no effect was found for Test face view, $F(1, 60) = 2.31$, $p = .14$, $\eta_p^2 = 0.07$, and no interaction was found between Type of Evaluation and Test face view, $F(2, 60) < 1$ (see Figure 4; see Table 6 for Hits and False Alarms, and Supplementary Table 4 for d' and RT).

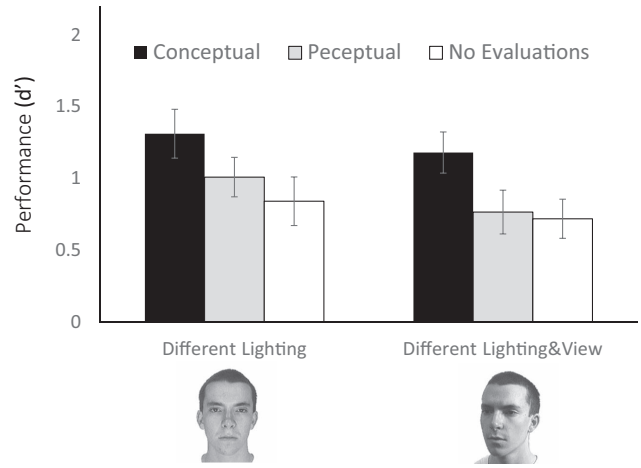

Figure 4. Recognition level (d') for faces that differed in lighting or lighting and view from the learned faces, following conceptual, perceptual, or no evaluations in Experiment 4. These images are used with permission by the Vision Science Lab.

Reaction time. A repeated measure ANOVA on RTs with Type of Evaluation (Conceptual, Perceptual, No Evaluations) and Test face view (different lighting, different lighting and view) as within-subject factors, showed no effect for Type of Evaluation, $F(2, 52) = 1.38$, $p > .25$, $\eta_p^2 = 0.05$, a marginally significant effect for Test face view, $F(1, 52) = 3.83$, $p = .061$, $\eta_p^2 = 0.13$, and a marginally significant interaction between Type of Evaluation and Test face view, $F(2, 52) = 2.47$, $p = .094$, $\eta_p^2 = 0.09$.

Learning phase. A paired-sample *t* test on RTs for the conceptual and perceptual evaluations revealed significantly shorter RTs for making conceptual than perceptual evaluations about faces, $t(30) = 3.37$, $p = .002$, Cohen's $d = 0.61$ (see Table 3), replicating the results of Experiment 1–3.

Discussion

Experiment 4 replicated our previous findings of better recognition for faces that undergo conceptual than perceptual evaluations. To determine whether this difference reflects a conceptual evaluation benefit or a perceptual evaluation cost, we compared performance level for faces that were learned with no evaluations. Results revealed better face recognition following learning faces with conceptual than no evaluations. Interestingly, there was no difference in recognition following perceptual evaluations than no evaluations, despite the fact that the former involved active evaluations of faces relative to passive viewing in the no evaluation condition. Whereas the benefit of conceptual over perceptual evaluations has been shown in several previous studies (e.g., Courtois & Mueller, 1979; Mueller et al., 1978; Winograd, 1976, 1981), to the best of our knowledge, this is the first time that the question of whether this effect is attributable to a conceptual evaluation benefit or a perceptual evaluation cost has been explicitly investigated. Our results importantly show that better recognition following conceptual than perceptual evaluations reflects a conceptual encoding benefit rather than a perceptual encoding cost. Interestingly, despite the fact that participants made an active evaluation of perceptual information during the learning phase, perceptual

Table 6
Mean, SD, and 95% Confidence Interval (CI) for Hits and False Alarm Rates for the Conceptual, Perceptual, and No Evaluations Conditions in Experiment 4

	Hit rate						False alarm rate	
	Conceptual evaluation		Perceptual evaluation		No evaluation			
Measure	Lighting	View and lighting	Lighting	View and lighting	Lighting	View and lighting	Lighting	View and lighting
Mean	.71	.66	.61	.52	.55	.50	.27	.28
SD	.25	.22	.22	.20	.25	.24	.17	.21
95% CI−	.62	.58	.53	.45	.46	.42	.21	.20
95% CI+	.80	.74	.69	.59	.64	.59	.33	.35

evaluations did not enhance recognition relative to passive viewing. These findings are consistent with previous findings showing that associating faces with person-related conceptual information during learning enhance face recognition. They add to previous findings by showing that nonconceptual evaluations, as the perceptual evaluations used in the current study, do not benefit face recognition. We discuss the implications of these findings in the general discussion.

Similar to results of Experiments 1 and 2, Experiment 4 also revealed shorter RT for conceptual than perceptual evaluations of faces during the learning phase. However, these responses were made after the offset of the face, during the presentation of the rating scale, rather than during the presentation of the face. This procedure was used to assure that faces are presented for the same exposure duration during learning across the different conditions. It is possible, however, that if we allow faces to remain on the screen during the evaluation period, the longer inspection time of additional features during conceptual evaluations would be reflected in longer RTs during conceptual than perceptual evaluations, as the *feature elaboration hypothesis* predicts (Winograd, 1981). To test this prediction, in Experiment 5 participants were asked to make conceptual and perceptual judgments during the presentation of the face that remained on the screen until response. Reaction times were measured from the onset of the face, and the face disappeared from the screen only after participants made their evaluation. Shorter RT to conceptual than perceptual evaluations during face presentation would provide a more direct evidence against the feature elaboration hypothesis (Winograd, 1976, 1981). It is noteworthy that in this experiment the length of exposure of each face was determined by the time it takes to make each type of evaluation. If conceptual evaluations are still faster, then conceptually learned faces would be presented for shorter durations during learning. Thus, in this experiment we can assess whether the conceptual evaluation benefit may be found even if faces are presented for shorter durations relative to perceptually learned faces.

Experiment 5

Method

Participants. Thirty-two participants (17 females, between the ages 22–44, mean age = 33.22, *SD* = 6.87) participated in this experiment. The experiment was conducted online on a survey platform named "Panel4all," similar to Amazon's "Mechanical Turk." All participants received payment for participating in the experiment ($4 for a 30-min experiment). The study was approved by the ethics committee of Tel Aviv University, and all participants gave informed consent to participate in the study.

Procedure. The stimuli and procedure were similar to Experiment 4 except the following changes: During the learning phase, following the presentation of the conceptual/perceptual/No Evaluations a face was presented with a scale beneath it and participants were asked to rate the face in the conceptual or perceptual conditions or passively view it in the no evaluations condition. The face was presented on the screen until participants made their decision in the conceptual and perceptual evaluations conditions and, similar to Experiment 4, for 1 s in the no evaluation condition. The recognition phase was identical to Experiment 4.

Results

Learning phase. We first report RT during the learning phase for the conceptual and perceptual evaluations. Paired *t* tests show a significantly shorter RT for conceptual than perceptual evaluations, $t(31) = 5.45$, $p < .0001$, Cohen's $d = 0.94$ (see Table 3 and Figure 5, left). Because faces were presented on the screen until participants made their decision, exposure duration for the faces presented following conceptual questions was significantly shorter (~500 ms on average) than for faces that were presented following perceptual questions. This difference should give an advantage for faces presented with perceptual questions during the recognition phase. We therefore examined whether recognition level was better following perceptual questions due to longer exposure duration during learning, or the conceptual processing was still more beneficial for face recognition despite the fact that the faces were presented for shorter exposure duration.

Recognition phase. A repeated measure ANOVA on the sensitivity score (d') with Type of Evaluation (Conceptual, Perceptual, No Evaluations) and Test face view (different lighting, different lighting & view) as within-subject factors revealed a main effect for Type of Evaluation ($F(2, 62) = 5.62$, $p = .006$, $\eta_p^2 = 0.15$). Post hoc Tukey's test revealed better recognition of the conceptually evaluated faces than the perceptually evaluated faces ($p = .04$, Cohen's $d = 0.48$), and faces presented with no evaluations ($p = .006$, Cohen's $d = 0.58$). Recognition was better for faces that differed from the learned images only in lighting as compared to faces that differed in both lighting and view, $F(1, 31) = 11.72$, $p = .002$, $\eta_p^2 = 0.27$, and no interaction was found between Type of Evaluation and Test face view, $F(2, 62) < 1$. (see

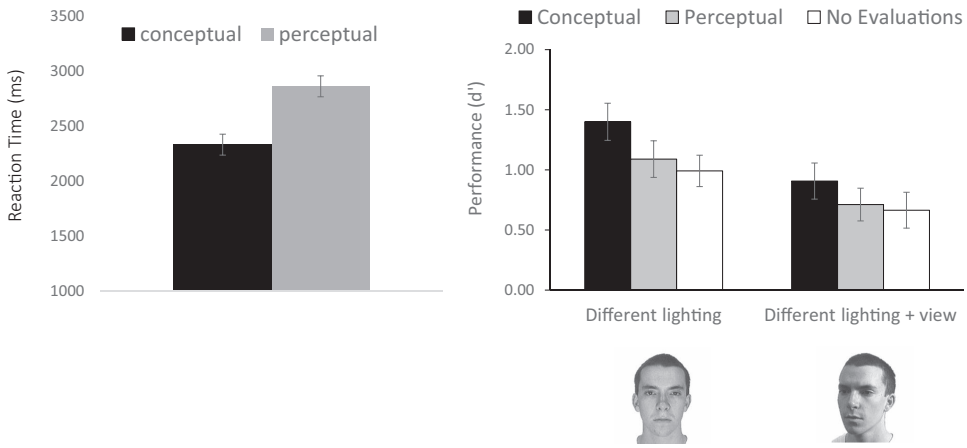

Figure 5. Left: Faster RT for conceptual than perceptual evaluations during the learning phase. Right: Recognition level (d') for faces that differed in lighting or lighting and view from the learned faces, following conceptual, perceptual, or no evaluations. These images are used with permission by the Vision Science Lab.

Table 7 for Hits and False Alarms and Supplementary Table 5 for d' and RT). This conceptual encoding benefit replicates findings of Experiment 4 relative to the perceptual and control conditions and Experiment 1–2 relative to the perceptual condition. Thus, despite longer exposure duration due to longer processing of faces during perceptual encoding, faces were better recognized following conceptual encoding (see Figure 5).

Discussion

According to the feature elaboration hypothesis (Winograd, 1981), face recognition is better following conceptual than perceptual evaluations because conceptual evaluations lead to more elaborated processing of facial features than conceptual evaluations. This hypothesis therefore predicts that processing time of faces during conceptual evaluations would be longer than during perceptual evaluations. In our previous experiments we found faster conceptual than perceptual evaluations, but these responses were made for faces that were presented on the screen for a fixed duration, after the offset of the face. In Experiment 5, participants made the evaluations during the presentation of the face that remained on the screen until a response was made, therefore allowing the needed inspection time for making a decision. Our findings show that evaluations were still faster for conceptual than perceptual encoding, providing no support for more elaborated encoding of a larger number of features during conceptual evaluations. Moreover, despite shorter exposure duration to conceptually evaluated faces, recognition was still better for conceptually than perceptually evaluated faces (Figure 5, right).

The *feature elaboration hypothesis* was proposed by Winograd (1981) based on a series of experiments in which participants were asked to search the most distinctive feature in a face or make an evaluation on a personality trait. The distinctive feature search task was chosen by Winograd to encourage the processing of more features, which based on the feature elaboration hypothesis, is expected to abolish the benefit of conceptual/trait evaluations. Results showed that when participants search for a distinctive feature, recognition was no longer better than when making trait evaluations. Whereas these findings were interpreted as a support for the feature elaboration hypothesis, we suggest that this inference is flawed for the following reasons: First, Winograd's choice of a distinctive feature search task was based on the assumption that to decide whether a feature is distinctive, participants compare the different facial features of a given face. According to Winograd "It is assumed that this task requires an extensive scan of the face, since feature distinctiveness is defined relative to the other features for a particular face. Thus, an elaborate encoding should be the outcome" (Winograd (1981) p. 182). However, distinctive features are evaluated not among different facial features of the same face

Table 7
Mean, SD, and 95% Confidence Interval (CI) for Hits and False Alarm Rates for the Conceptual, Perceptual, and No Evaluations Conditions in Experiment 5

	Hit rate						False alarm rate	
	Conceptual evaluation		Perceptual evaluation		No evaluation			
Measure	Lighting	View and lighting	Lighting	View and lighting	Lighting	View and lighting	Lighting	View and lighting
Mean	.76	.62	.66	.55	.62	.54	.30	.32
SD	.19	.20	.20	.23	.28	.22	.22	.24
95% CI−	.69	.55	.59	.47	.52	.46	.22	.24
95% CI+	.83	.69	.73	.63	.72	.61	.37	.40

(e.g., the face has large nose not relative to an averaged eye shape, hair style or lip thickness), but relative to the same feature of other identities (e.g., the face has a much larger nose relative to noses of other faces; Rhodes, Brennan, & Carey, 1987; Valentine, 1991). Thus, a search for a distinctive feature is not necessarily based on scanning of all facial features, but on a comparison to other identities in memory. Furthermore, a distinctive feature is typically salient (e.g., a very distinct hair style, very thick eyebrows) and therefore can attract attention, which may result in the opposite effect of reduced feature scanning of the rest of the facial features. Thus, the basic premise of using this manipulation as a way to encourage elaborated feature scanning is unfounded. Second, the fact that searching a distinctive feature abolishes the difference between perceptual and conceptual evaluations does not mean that the same type of processing takes place in both conditions. Different types of processing styles may lead to similar performance level in a recognition task. Third, a distinctive feature search task may in fact encourage a conceptual rather than perceptual processing. When a face has a distinctive feature, we may encode this feature verbally (e.g., "the face with the large nose"). When we see a face with a large nose at test, we may retrieve this conceptual label and the face associated with it, which may lead to better recognition. Fourth, improved performance for the distinctive feature search may be based on the fact that distinctive faces are more memorable rather than on the scanning itself. This was in fact tested by Winograd (1981) in Experiment 3, in which he compared directly the effect of evaluation of a distinctive versus a nondistinctive feature evaluation and a distinctive feature search. Whereas mean hit rate was numerically better for the distinctive feature scan condition (.81) than the distinctive feature evaluation condition (.74), mean FA rates was numerically lower for distinctive feature evaluations (.14) than the distinctive feature search (.17), suggesting that an accuracy score that takes into account both Hits and FAs would be more similar between distinctive feature evaluation and distinctive feature search (Hits-FAs = .60 and .64, respectively) than Hit rate alone. Winograd indicated that because FA rates of the individual subjects were lost, they could not include them in their statistical analysis, which were performed only on Hit rates, therefore probably overestimating the effect of distinctive feature scanning relative to distinctive feature evaluations. Furthermore, mean hit rates for the nondistinctive feature evaluation was numerically lower (.69) and FA rates were higher (.18) than for the distinctive feature evaluation, indicating the judging a distinctive feature contributes to face recognition regardless of the feature scan procedure. The ANOVA that was performed on hit rates of the three conditions revealed a significant main effect of condition but was not followed by post hoc tests to reveal which of the conditions were significantly different. It is likely that the main effect of condition reflects the lower performance on the nondistinctive feature evaluations, relative to the distinctive feature evaluations and the distinctive feature search than the latter two conditions. Finally, the most direct measure of the feature elaboration hypothesis, which is how long it takes to perform these evaluations/feature scans during learning, was not measured. The tests were given to a large group of subjects during an introductory psychology class. The faces were presented on a screen to all subjects and the participants made the feature or trait evaluations as well as the recognition task by writing their answers in a booklet. Taken together, the feature elaboration hypothesis was based on unsupported assumptions, confounds of the effects of feature scanning and feature distinctiveness on recognition, incomplete data analyses and inadequate experimental procedures. Our findings across five experiments show that conceptual evaluations are faster than perceptual evaluations and therefore provide no support for the feature elaboration hypothesis.

General Discussion

Across four face recognition experiments, we repeatedly showed better recognition for faces that underwent conceptual than perceptual encoding. This benefit was not attributable to longer processing time for faces that underwent conceptual evaluations during learning, as RTs were significantly shorter for conceptual than perceptual evaluations across all experiments (see Table 3). These findings therefore indicate that conceptual encoding does not involve more elaborated processing of facial information (i.e., the feature elaboration hypothesis, Winograd (1981). Instead, faster conceptual than perceptual evaluations suggest that it is more efficient to process faces as concepts rather than percepts. Second, the benefit of conceptual relative to perceptual evaluations in face recognition was found for both local and global feature evaluations (Experiments 1 and 2), excluding the possibility that the conceptual evaluation benefit merely reflects the processing of global rather than local facial features. Third, the *conceptual encoding benefit* cannot be explained by larger variability in trait than perceptual evaluations across faces, which may increase the perceptual distance among them during learning, as the variability of conceptual rating was not larger than perceptual rating (Experiment 3). Fourth, we found no evidence that the conceptual evaluation benefit is due to greater redundancy in perceptual than conceptual questions, which may lead to more limited encoding of facial information. Finally, our final experiments importantly show that better recognition for conceptually than perceptually encoded faces is not because perceptual evaluations interfere with face recognition, but because conceptual evaluations enhance face recognition (Experiment 4 & 5). To the best of our knowledge, this is the first evidence that better recognition for faces that underwent conceptual than perceptual evaluations reflects a *conceptual evaluation benefit* rather than a perceptual evaluation cost. Our findings also show that active perceptual evaluations did not improve face recognition more than learning faces passively. These results imply that not all face-related information processing benefits face recognition. Face recognition benefits from the processing of conceptual but not from perceptual face-related information. In fact, this lack of benefit during a condition in which participants are more engaged than passively viewing faces may imply that attending perceptual information may involve some cost.

Our studies also revealed that the *conceptual encoding benefit* in face recognition is generalized to other images of the learned identities that differ in view and lighting. An important indication that a stimulus in general, and a face in particular, was processed conceptually is the finding that the conceptual benefit is not image-specific but improves recognition of the identity of the learned face. Previous studies that showed better recognition following trait inferences than perceptual evaluations used the same face image in study and test (Bower & Karlin, 1974; Courtois & Mueller, 1979; Mueller et al., 1978; Strnad & Mueller, 1977; Winograd, 1976, 1981). The generalization of the effect to new

images of the same identity is especially important in light of many studies that indicate that same-image recognition may reflect low-level pictorial recognition rather than recognition of person identity (Bindemann & Sandford, 2011; Bruce, 1982; Burton, Wilson, Cowan, & Bruce, 1999; Longmore, Liu, & Young, 2008; Megreya & Burton, 2006; White, Burton, Jenkins, & Kemp, 2014). Thus, our results indicate that the *conceptual encoding benefit* in face recognition reflects person recognition rather than image-based face recognition.

Our findings that face recognition benefits from conceptual than perceptual encoding are consistent with previous reports that face recognition benefits from the association of person-related information during learning (Tanaka & Pierce, 2009; Yovel et al., 2012; Schwartz & Yovel, 2016). Although the two paradigms are different, they both involve the association of faces with conceptual person-related information (name, occupation, traits) and therefore encourage the representation of faces as concepts rather than percepts. Whereas faces have been typically studied as perceptual stimuli, and most studies of face recognition have examined different types of perceptual manipulations that may influence their representations (Maurer, Grand, & Mondloch, 2002; McKone & Yovel, 2009; McWeeny, Young, Hay, & Ellis, 1987; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 2013), here we highlight the fact that face recognition is not a purely perceptual process but significantly benefits from the generation of a conceptual representation. This is also reflected in the faster RTs for conceptual than perceptual evaluations during the learning stage.

Whereas most studies of face recognition have focused on the role of perceptual information in face recognition (Burton et al., 2005, 2011; Burton, Kramer, Ritchie, & Jenkins, 2016; Jenkins & Burton, 2011; Ritchie & Burton, 2017), other studies have also shown an important role for social factors in face recognition (Bernstein, Young, & Hugenberg, 2007; Hugenberg, Wilson, See, & Young, 2013; Rule, Ambady, Adams, & Macrae, 2007). Social processing of faces has been primarily studied with respect to facial expressions, traits, race or gender. Nevertheless, the identity of a person is highly important for intact social interaction and indeed studies have shown that social factors influence face recognition. For example, Wilson, See, Bernstein, Hugenberg, and Chartier (2014) have shown that expectation for future interaction with people, improves face recognition for out-group members (Wilson et al., 2014). A study by Van Bavel and Cunningham (2012) found that social identification with a group improves face recognition for members of that group both for own and other group members (Van Bavel & Cunningham, 2012). To account for these and other similar findings, Hugenberg and colleagues suggested that social factors may motivate individuation of own-group members and thus increase the dissimilarity between their representation and improve recognition of different individuals (Hugenberg et al., 2013). This suggestion is consistent also with reports that show improved recognition for faces of other races that are associated with different names but not for faces that are associated with the same label (Gordon & Tanaka, 2011; Tanaka & Pierce, 2009). Our recent study, however, further showed that individuation that is not associated with person-related relevant information, does not improve face recognition (Schwartz & Yovel, 2016). In particular, associating different faces with different person-unrelated labels (e.g., symbols, object names) did not improve face recognition relative to unlabeled faces. In contrast, associating faces with person-related labels, such as person names and occupations, significantly improves face recognition relative to nonlabeled faces. We therefore suggested that individuation alone may not account for improved face recognition, and instead highlight the benefit of associating person-related semantic/social information during learning to face recognition to generate a conceptual representation.

The effect of semantic labels on visual processing has been shown in several previous studies. For example, Lupyan and colleagues have shown that adding a word label to a visual stimulus, influences its representation, which consequently enhance or interfere with performance on a visual task (e.g., Lupyan, 2008; Lupyan & Spivey, 2008). Doyle and Lindquist (2017) have shown that word labels associated with emotional facial actions during learning influence their representation during recognition in a way that is consistent with the learned label. These studies concluded that conceptual knowledge modifies the representation of the stimuli it is associated with, highlighting the importance of linking cognition and perception to understand recognition. Our findings are consistent with these studies in that they show that different labels (traits versus perceptual) associated with the same stimuli significantly modify the way they are encoded and then retrieved from memory.

Taken together, current and previous findings highlight the importance of encoding faces as concepts rather than as pure percepts for face recognition. In real life, the face recognition system is primarily engaged with familiar faces. We hardly ever need to match the identity of unfamiliar faces, but mostly match incoming faces to their representation in memory. Encoding of familiar faces is never purely perceptual and always involves conceptual/social information. Also, the representation of faces in memory is associated with conceptual/social information. Indeed, it is well established that our ability to recognize people out of context is inferior (e.g., Mandler, 1980). We therefore suggest that the benefit that was found in the current and previous studies for conceptually encoded faces implies that it is the conceptual processing of familiar faces rather than the perceptual exposure to unfamiliar faces that shapes the way faces are encoded, stored and retrieved from memory. This suggestion is consistent with previous findings that passive exposure over many years to newborn faces by neonatology nurses did not improve newborn face recognition, whereas training with a few face-name associations did improve performance on the same task (Yovel et al., 2012).[1] Therefore, to understand face recognition it is essential to understand not only their perceptual representation but also their conceptual representation, which provides them with social meaning for interpersonal interaction.

Our results are also relevant to studies that examined the effect of level of processing on recognition (Craik & Lockhart, 1972; Craik & Tulving, 1975). These studies were primarily performed on word stimuli, reporting better recognition for words that under-

[1] Cassia, Picozzi, Kuefner, and Casati (2009) revealed, similar to Yovel et al. (2012), no difference in performance between newborn faces in neonatology nurses and control participants, but a larger inversion effect for newborn faces in nurses. These findings suggest a dissociation between the absence of an effect of perceptual experience on recognition of upright faces and the appearance of an inversion effect, which should be further examined in future studies.

went semantic than visual evaluations during study. Whereas the extension of these studies to faces revealed similar findings (Bower & Karlin, 1974; Courtois & Mueller, 1979; Mueller et al., 1978; Strnad & Mueller, 1977; Winograd, 1976, 1981), unlike words, semantic and perceptual evaluations are not independent when faces or other visual stimuli are evaluated and both depend on physical features of the stimuli. A recent study compared the effect of semantic and perceptual rating on word and visual images (e.g., doors, watches, scenes) revealed overall smaller semantic benefit for visual stimuli than words (Baddeley & Hitch, 2017). It is noteworthy that the study used pleasantness rating for semantic encoding of doors and clocks. Such rating is hardly ever used for these types of objects, which are more likely to be represented in memory according to their functionality or esthetic features, and therefore is not likely to transform them to meaningful conceptual representations. Our findings are therefore consistent with the interpretation that the magnitude of the conceptual encoding benefit is linked to the type of semantic information that is associated with the stimulus in memory rather than whether it is visual or verbal (Baddeley & Hitch, 2017). We therefore suggest that the conceptual encoding benefit that we revealed for faces is not specific to face stimuli but is expected to be found for any stimulus that is processed based on its conceptual meaning rather than perceptual features. We further posit that such a conceptual encoding benefit may be the basis for the effects of expertise with nonface stimuli. Expertise for nonface stimuli has also been primarily studied from a perceptual perspective (Bukach, Gauthier, & Tarr, 2006; McKone, Kanwisher, & Duchaine, 2007), but it is typically accompanied by a rich vocabulary of labels and conceptual knowledge about objects of expertise. Future studies need to examine the role of conceptual information in expertise for any type of stimuli (e.g., wine, music, objects).

In conclusion, our findings repeatedly show that making conceptual evaluations about faces significantly improves person recognition. We further suggest that this conceptual benefit is because the representation of faces is based on the active social interaction that we have with familiar faces, rather than the passive viewing of unfamiliar faces. Familiar faces are never encoded or stored in memory as purely perceptual representations, but as meaningful social stimuli. This highlights an important role for social and conceptual processing in person recognition and therefore suggests that rather than primarily focusing on perceptual facial aspects in the study of face recognition, social and semantic factors should be also considered to obtain a full understanding of the various factors that contribute to person recognition.

References

Abudarham, N., & Yovel, G. (2016). Reverse engineering the face space: Discovering the critical features for face identification. *Journal of Vision, 16,* 40. http://dx.doi.org/10.1167/16.3.40

Baddeley, A. D. (1978). The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. *Psychological Review, 85,* 139–152. http://dx.doi.org/10.1037/0033-295X.85.3.139

Baddeley, A. D., & Hitch, G. J. (2017). Is the Levels of Processing effect language-limited? *Journal of Memory and Language, 92,* 1–13. http://dx.doi.org/10.1016/j.jml.2016.05.001

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18,* 706–712. http://dx.doi.org/10.1111/j.1467-9280.2007.01964.x

Bindemann, M., & Sandford, A. (2011). Me, myself, and I: Different recognition rates for three photo-IDs of the same person. *Perception, 40,* 625–627. http://dx.doi.org/10.1068/p7008

Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology, 103,* 751–757. http://dx.doi.org/10.1037/h0037190

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73,* 105–116. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7059746

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences, 10,* 159–166. http://dx.doi.org/10.1016/j.tics.2006.02.004

Burton, A. M., Jenkins, R., Hancock, P. J. B., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology, 51,* 256–284. http://dx.doi.org/10.1016/j.cogpsych.2005.06.003

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology, 102,* 943–958. http://dx.doi.org/10.1111/j.2044-8295.2011.02039.x

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40,* 202–223. http://dx.doi.org/10.1111/cogs.12231

Burton, M. A. (2013). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology, 66,* 1467–1485. http://dx.doi.org/10.1080/17470218.2013.800125

Burton, M. A., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10,* 243–248. http://dx.doi.org/10.1111/1467-9280.00144

Cassia, V. M., Picozzi, M., Kuefner, D., & Casati, M. (2009). Short article: Why mix-ups don't happen in the nursery: Evidence for an experience-based interpretation of the other-age effect. *Quarterly Journal of Experimental Psychology, 62,* 1099–1107.

Coin, C., & Tiberghien, G. (1997). Encoding activity and face recognition. *Memory, 5,* 545–568. http://dx.doi.org/10.1080/741941479

Courtois, M. R., & Mueller, J. H. (1979). Processing multiple physical features in facial recognition. *Bulletin of the Psychonomic Society, 14,* 74–76. http://dx.doi.org/10.3758/BF03329404

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11,* 671–684. http://dx.doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104,* 268–294. http://dx.doi.org/10.1037/0096-3445.104.3.268

Doyle, C. M., & Lindquist, K. A. (2017). Language and emotion: Hypotheses on the constructed nature of emotion perception. In J. M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (pp. 415–432). New York, NY: Oxford University Press.

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia, 44,* 576–585. http://dx.doi.org/10.1016/j.neuropsychologia.2005.07.001

Eysenck, M. W. (1978). Levels of processing: A critique. *British Journal of Psychology, 69,* 157–169. http://dx.doi.org/10.1111/j.2044-8295.1978.tb01643.x

Gordon, I., & Tanaka, J. W. (2011). The role of name labels in the formation of face representations in event-related potentials. *British Journal of Psychology, 102,* 884–898. http://dx.doi.org/10.1111/j.2044-8295.2011.02064.x

Hugenberg, K., Wilson, J. P., See, P. E., & Young, S. G. (2013). Towards a synthetic model of own group biases in face memory. *Visual Cognition, 21,* 1392–1417. http://dx.doi.org/10.1080/13506285.2013.821429

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 366,* 1671–1683. http://dx.doi.org/10.1098/rstb.2010.0379

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 77–100. http://dx.doi.org/10.1037/0096-1523.34.1.77

Lupyan, G. (2008). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General, 137,* 348.

Lupyan, G., & Spivey, M. J. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology, 18,* R410–R412.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Science, 6,* 255–260.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11,* 8–15. http://dx.doi.org/10.1016/j.tics.2006.11.002

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review, 16,* 778–797. http://dx.doi.org/10.3758/PBR.16.5.778

McWeeny, K. H., Young, A. W., Hay, D. C., & Ellis, A. W. (1987). Putting names to faces. *British Journal of Psychology, 78,* 143–149. http://dx.doi.org/10.1111/j.2044-8295.1987.tb02235.x

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition, 34,* 865–876. http://dx.doi.org/10.3758/BF03193433

Mueller, J. H., Carlomusto, M., & Goldstein, A. G. (1978). Orienting task and study time in facial recognition. *Bulletin of the Psychonomic Society, 11,* 313–316. http://dx.doi.org/10.3758/BF03336841

Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning & Verbal Behavior, 16,* 151–171. http://dx.doi.org/10.1016/S0022-5371(77)80044-3

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

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology, 19,* 473–497. http://dx.doi.org/10.1016/0010-0285(87)90016-8

Ritchie, K. L., & Burton, M. A. (2017). Learning faces from variability. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 70,* 897–905.

Rule, N. O., Ambady, N., Adams, R. B., Jr., & Macrae, C. N. (2007). Us and them: Memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin & Review, 14,* 687–692. http://dx.doi.org/10.3758/BF03196822

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22,* 36–71. http://dx.doi.org/10.1016/0010-0285(90)90003-M

Schwartz, L., & Yovel, G. (2016). The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General, 145,* 1493–1511. http://dx.doi.org/10.1037/xge0000220

Strnad, B. N., & Mueller, J. H. (1977). Levels of processing in facial recognition memory. *Bulletin of the Psychonomic Society, 9,* 17–18. http://dx.doi.org/10.3758/BF03336915

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 46,* 225–245. http://dx.doi.org/10.1080/14640749308401045

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective & Behavioral Neuroscience, 9,* 122–131. http://dx.doi.org/10.3758/CABN.9.1.122

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 43,* 161–204. http://dx.doi.org/10.1080/14640749108400966

Van Bavel, J. J., & Cunningham, W. A. (2012). A social identity approach to person memory: Group membership, collective identification, and social role shape attention and memory. *Personality and Social Psychology Bulletin, 38,* 1566–1578. http://dx.doi.org/10.1177/0146167212455829

White, D., Burton, A. M., Jenkins, R., & Kemp, R. I. (2014). Redesigning photo-ID to improve unfamiliar face matching performance. *Journal of Experimental Psychology: Applied, 20,* 166–173. http://dx.doi.org/10.1037/xap0000009

Wilson, J. P., See, P. E., Bernstein, M. J., Hugenberg, K., & Chartier, C. (2014). Differences in anticipated interaction drive own group biases in face memory. *PLoS ONE, 9,* e90668. Advance online publication. http://dx.doi.org/10.1371/journal.pone.0090668

Winograd, E. (1976). Recognition memory for faces following nine different judgments. *Bulletin of the Psychonomic Society, 8,* 419–421. http://dx.doi.org/10.3758/BF03335185

Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 181–190. http://dx.doi.org/10.1037/0278-7393.7.3.181

Young, A. W., Hellawell, D., & Hay, D. C. (2013). Configurational information in face perception. *Perception, 42,* 1166–1178. http://dx.doi.org/10.1068/p160747n

Yovel, G., Halsband, K., Pelleg, M., Farkash, N., Gal, B., & Goshen-Gottstein, Y. (2012). Can massive but passive exposure to faces contribute to face recognition abilities? *Journal of Experimental Psychology: Human Perception and Performance, 38,* 285–289. http://dx.doi.org/10.1037/a0027077

Received April 23, 2018
Revision received September 27, 2018
Accepted September 27, 2018 ■