THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
58A (2), 235–250

Psychology Press
Taylor & Francis Group

When context hinders! Learn–test compatibility in face recognition

Helmut Leder and Claus-Christian Carbon
Freie Universität Berlin, Berlin, Germany

Some theories of holistic face processing propose that parts in faces (eyes, nose, mouth, etc.) are not explicitly represented. So far, the empirical evidence has shown that whole-to-part superiority is found when wholes are learned. We substantiated this using photographic faces. More importantly, we investigated whether learning parts also reveals holistic effects. This has not been attempted before. Four experiments showed that after learning facial parts, recognition of these parts was disrupted when the part was shown in the full face. This distraction effect was strongest when perceivers were not directed to focus on a particular facial feature. Thus, it is very difficult to ignore irrelevant parts in faces. In fact, this might be the essence of holistic face processing.

In face recognition, holistic processing has been defined as a representation in which parts are not explicitly represented (Tanaka & Farah, 1993). This holistic processing proposal has been supported by findings that recognition of facial parts from facial context (full faces) is superior to recognition of facial parts presented alone. Tanaka and Farah had participants learn schematic faces composed of shared features. In a recognition test, participants had to choose which face corresponded to a learned face from either two facial parts (e.g., two different noses) or two faces differing only in these parts. Decisions were more difficult when the parts were presented in isolation. Tanaka and Farah suggest that this kind of holistic superiority is specific to faces presented upright, since turning the faces upside-down reduced the effect. Moreover, Tanaka and Farah did not find similar effects using schematic houses or scrambled faces. Nevertheless, similar effects of whole-to-part superiority at test were found by Donnelly and Davidoff (1999) with faces, schematic houses, and very simple house drawings. Tanaka and Gauthier (1997) also found similar effects with so-called greeble experts who had learned to distinguish face-like artificial objects.

Correspondence should be addressed to Helmut Leder, Department of Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. Email: leder@experimental-psychology.de

This research was supported by a grant to Leder from the Deutsche Forschungsgemeinschaft (DFG, Le1286). We particularly thank G. Harsányi and all students for their help in conducting the experiments. We thank Eric Kearney and Steve Pawlett for proof-reading the manuscript. Moreover, we thank Mike Burton, Jules Davidoff, and an anonymous reviewer for their helpful comments.

http://www.tandf.co.uk/journals/pp/02724987.html DOI:10.1080/02724980343000936

None of these studies tested whether part recognition following part learning is also affected by the context of a full face at test. One study investigated a related question. Farah, Tanaka, and Drain (1995) presented faces in the learning phase as an ensemble of four different features. When tested as wholes these faces did not show the usual whole-to-part superiority. However, the features tested in this experiment were shared, and a distinctive part condition was not employed in the study phase. It is an important question to address whether part–part as compared to part–whole tests also reveal aspects of holistic processing.

An alternative way of defining holistic processing stems from Farah, Wilson, Drain, and Tanaka (1998). They inferred holistic processing from interference effects. When a part of a face had to be matched in a simultaneous matching task, ignoring irrelevant but compatible parts was difficult.

In the present experiments we investigated a similar effect in that we had people learn particular facial parts and tested whether embedding these parts in whole faces also produced interference effects. There are three alternative hypotheses. If faces are processed in a strict holistic way, in that part processing is always superior from wholes than from isolated features, then we expected whole-to-part superiority in all conditions of our experiments. Alternatively, if learning parts imposes a strict part-based representation we expect that an additional context at test might be ignored, and whole and part conditions at test might be equally good. Finally, if holistic processing means that part processing is disrupted from full faces due to the difficulty of ignoring irrelevant information, then after learning parts an unfamiliar full face at test should have a strong interference effect.

In contrast to previous studies, we did not only test whole–whole and whole–part conditions. We also compared experiments in which it was easier to focus on a specific face region. In Experiments 2 and 4 the specific facial parts consisted of eye regions only, thus allowing subjects to focus on that region in the test-phase. Conversely, in Experiments 1 and 3 component parts consisted of eyes, noses, or mouths, making focusing more difficult. Whether focusing subjects' attention on a particular feature affects interference between wholes and parts will be revealed by cross-experimental comparisons.

In four experiments a recognition task similar to the one developed by Leder and Bruce (2000) was used, in which participants first learned a set of faces and then were tested by selecting the appropriate name for each face presented in the recognition phase.

In all our experiments we used both upright and upside-down orientations at test. Inversion has been shown to disrupt holistic processing (Tanaka & Farah, 1993) as well as configural processing (Leder & Bruce, 2000). If configural processing is involved in some conditions more than in others, we expect differential inversion deficits in these conditions. If, however, no specific disruption is found, this has to be interpreted as a lack of a condition-specific configural or holistic processing.

Previous studies where it was shown that it was difficult to recognize parts when they were embedded in whole faces used shared features between the different faces that had to be recognized. Moreover, only schematic faces have been used (Donnelly & Davidoff, 1999; Tanaka & Farah, 1993). In real faces, eyes, noses, and mouths are individually distinctive, and real faces contain texture. In our experiments we used artificial computer-generated faces, which differed from each other in terms of eyes, nose, and mouth (Experiment 1) or eye region only (Experiment 2). In later experiments, we used real faces, which were tested in component and full-face recognition of eyes, nose, and mouth (Experiment 3) or eyes only (Experiment 4).

In these experiments we familiarized faces in a number of learning trials as in Leder and Bruce (2000). Nevertheless, the results might reveal strategies that are typical for the processing of rather new and unfamiliar faces. Young, Hay, McWeeny, Flude, and Ellis (1985) have shown that processing of unfamiliar faces relies more on outer features. Since the outer features are available in Experiments 3 and 4, but not in Experiments 1 and 2, we may expect subjects to use processes relying on gross outer features such as hairstyle in the former experiments. However, individual features, such as those learned in Experiments 1 and 2, are very similar to each other, and so their coding must presumably rely on distinctive parts. We examine evidence for differential processing in patterns of inversion effects across the experiments.

EXPERIMENT 1

Method

Participants

A total of 16 graduate students and undergraduates (6 females, overall mean age 25.3 years) from the Freie Universität Berlin participated for course credit. All participants were tested individually.

Materials

Features from six digital photographs of young male persons from the ATR face databank (Leder, 1998) were used to produce the stimuli for Experiment 1. First, an artificial face was created, which consisted of carefully manipulated eyes, nose, and mouth from three different faces within a new facial outline. Based on this face, six different stimuli were created. Component features, eyes, mouths, and noses were carefully copied into the context of the "basic face". Six versions were created, with each having only one critical facial part. Two faces had distinctive, individual eye regions, two had specific noses, and two had specific mouths. Figure 1 shows all six faces used in the "full" face conditions of Experiment 1. Thus each face had one critical part plus parts shared by four of the other faces. Four versions of each face were used: one full-face upright, one full-face upside-down, one part-face upright, and one part-face upside-down. To create part versions, the individually distinct feature of each face was cut out and used in isolation (see Figure 3), again either upright or upside-down.

Twelve German names were selected and randomly assigned to one of two sets. The names were Jürgen, Alfred, Gregor, Robert, Hannes, Dieter (Set 1), and Arnold, Konrad, Justus, Tobias, Reiner, Sascha (Set 2).

Procedure and design

Participants were tested individually. The order of trials within each block was randomized by an experimental program.

Study phase

At the beginning of each experimental session, the participants were told they would be exposed to six male persons' faces, which they should try to memorize and later recognize. In the learn–full learn-

Figure 1. Stimuli used in Experiment 1. Faces differ in respect to one distinctive feature from each other while all other features are shared with four other faces. Distinctive eyes (Faces 2 and 6), mouth (Faces 1 and 5), nose (Faces 3 and 4).

ing condition they were told to view six faces, in the isolated condition (learn–part) they were told to memorize six facial parts belonging to six different persons. All material was presented upright. The learning condition was manipulated between subjects (8 participants each). The study phase consisted of three blocks in which each of the six stimuli (either full faces or part versions) were presented on the screen for five seconds along with a short sentence stating "This is . . ." plus the name.[1]

Within each of these blocks, the stimuli were presented in a randomized order. A block in which each face was shown once followed the learning block, and the participants' task was to tell the experimenter the name of each face. During this phase, each face was shown for five seconds along with the question "Who is this?" After three seconds, the correct name was shown beneath the stimulus to provide feedback. If participants scored fewer than five faces correct in this block, blocks of two additional learning trials were repeated until each participant met this criterion.

Test phase

After a short break the test block began. All six names were shown along with a number beneath ranging from 1 to 6, indicating which number to press on the keyboard for each name. In each trial one test face was presented beneath the list of names. Participants were instructed to press the number corresponding to the stimulus person's name. The test stimulus then disappeared, and subjects proceeded to the next stimulus, self-paced, by pressing the space bar. Each face was shown four times, either whole (test–full) or part (test–part), and in two orientations (upright and inverted), thus yielding a total of 24 trials in the test for each version. The order of the stimuli in the test was randomized for each participant.

---

[1] The German versions were used in all experiments.

Results and discussion

Table 1 shows the mean proportion of correctly identified stimuli in the test phase, separately for those participants who learned the full face and for those who learned the isolated parts.

Data were submitted to a three-way analysis of variance (ANOVA) with learn size condition (learn–full vs. learn–part) as between-subjects factor and test size (test–full and test–part) and orientation (upright and inverted) as within-subjects factors. The analyses revealed an almost significant main effect of learn size, $F(1, 14) = 3.907$, $p = .068$, *ns*, with percentage of correct rates being higher after the learning of parts (after learning parts, .656; after learning full faces, .531). A significant main effect of orientation was obtained, $F(1, 14) = 16.579$, $p < .01$, with upright being easier than upside-down (upright, .672; inverted, .516), and a significant effect of test size, part versus whole, $F(1, 14) = 12.411$, $p < .01$ (recognition of parts, .682; recognition of full faces, .505). The interaction between orientation and test size, $F(1, 14) = 6.054$, $p < .05$, was also significant, as well as the three-way interaction between orientation, test size, and learn size, $F(1, 14) = 9.459$, $p < .01$. Moreover, there was an interaction between test size and learn size, $F(1, 14) = 9.663$, $p < .01$, which is illustrated (for the upright trials only) in Figure 2.

Most interesting are the results concerning the whole-to-part superiority at test. When full faces had been learned, and only upright faces are considered, there was a tendency towards the expected difference (part, .542, vs. full, .708), but the effect was surprisingly weak. We ran a one-tailed *t* test, in order to test the hypothesis of whole-to-part superiority, using only upright conditions. The test revealed a $t(7) = 1.871$, $p = .052$, *ns*. The interaction as well as the main effect of whole-to-part comparisons in Experiment 1 was caused by the results in the part learn conditions. When isolated parts were learned, there was no superiority of a full-face context in the test but facial parts were recognized significantly better. This was revealed in an undirected, two-tailed *t* test with $t(7) = 4.123$, $p < .01$, as we had no directional hypothesis for that comparison although part–part superiority would be in accordance with Tulving and Thomson's (1973) encoding specificity. While the former result is in accordance with the expectation of the holistic processing hypothesis, the latter result stands in contrast to the prediction of a general superiority for full faces. The new whole after learning parts reveals a strong impact of facial features on the recognition of parts. In terms of face processing this result supports the idea that facial parts are difficult to isolate from the facial whole. The finding thus provides evidence for a kind of holistic representation in which access to a critical part is hindered by task-irrelevant information.

TABLE 1
Mean identification rates of Experiment 1

	Test							
	Test–full–up		*Test–part–up*		*Test–full–in*		*Test–part–in*	
Learn	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Full	.708	.17	.542	.15	.333	.18	.542	.26
Part	.542	.25	.896	.09	.438	.20	.750	.18

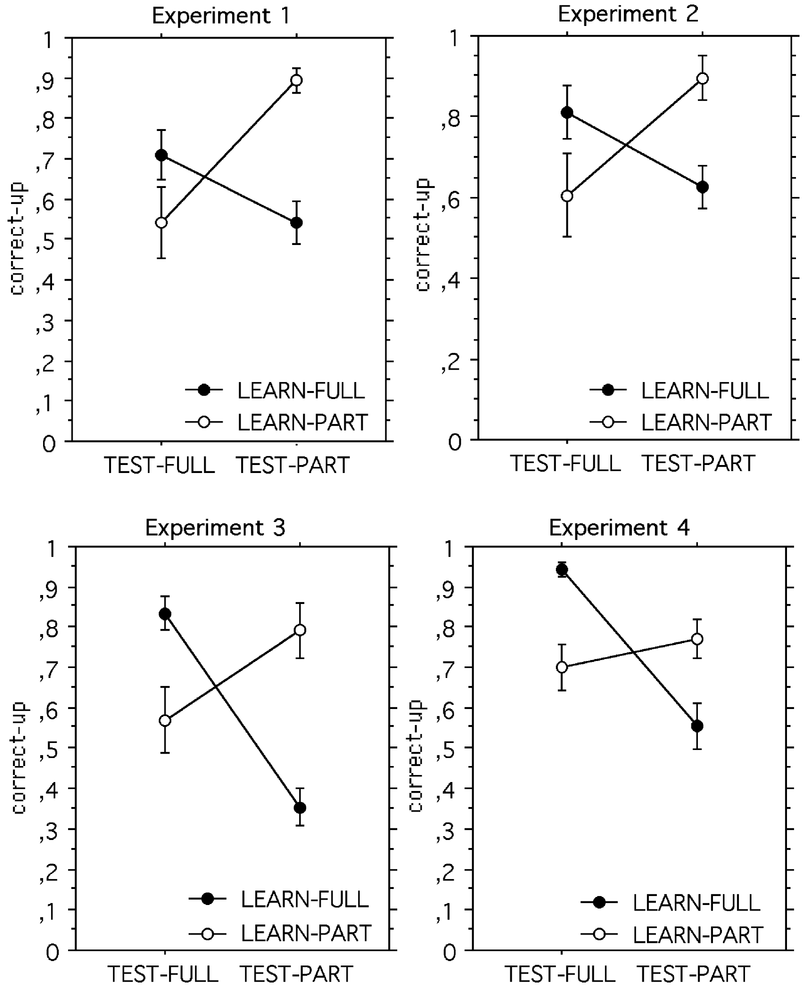

Figure 2. Interactions between learn and test conditions (for upright conditions only) in all four experiments.

The size of the inversion effects also depended on the conditions during the learning *and* testing phases. The full faces reveal particularly large inversion deficits when the full faces had been learned (inversion effect, i.e., upright minus inverted, .375), while full faces after learning of parts showed significant smaller inversion deficits (inversion effect, .104). The difference between both conditions (comparing full faces at test after full respectively part-based faces had been learned) was found to be significant by a two-tailed t test with $t(14) = 2.244$, $p < .05$. These results are in accordance with the hypothesis that faces learned as wholes are more strongly affected by inversion (Leder & Bruce, 1998) as they probably contain a number of important configurations. Nevertheless, both inversion effects are significantly larger than zero, one-tailed t tests, $ts(7) > 1.9$, $ps < .05$, and therefore reveal that the parts—eyes, noses, and mouth—probably contain elements of local relational configuration that are sensitive to inversion. Alternatively, parts after being learned in isolation might interact with the context of the full face. In this case the hindering effect of the full face is due to configurations from the whole that prevent successful parsing. This interference effect is similar to the Young, Hellawell, and Hay (1987) composite effect.

However, there were no inversion effects when participants learned full faces and had to recognize isolated parts. Inversion deficits are explained by at least two competing hypotheses: Either the full faces are processed holistically or they contain a number of helpful configurations, which are difficult to process when faces are turned upside-down. Both hypotheses are compatible with the results of Experiment 1.

To investigate whether the whole inevitably distracts from successfully parsing the critical region, we used different faces in Experiment 2, which differed from each other only in respect to the eye region. Thus, in the part versions of Experiment 2 only different eye regions were used to allow participants to focus on a specific face region in all conditions. If holistic processing is revealed by interference of irrelevant information then interference effects should also be found here. However, if the effect of Experiment 1 is completely caused by the difficulty of focusing on face internal parts, then no interference effects might be expected in the part–whole conditions of Experiment 2.

EXPERIMENT 2

Method

Participants

A total of 16 graduate students and undergraduates (12 females, overall mean age 25.0 years) from the Freie Universität Berlin participated for course credit. All participants were tested individually.

Materials

The material used in Experiment 2 was very similar to that of Experiment 1 except that the faces differed only in respect to eye regions. Thus, six distinct eye regions were embedded in the full faces. The different eye regions are shown in Figure 3. All other features were held constant. The same name sets as those in Experiment 1 were used.

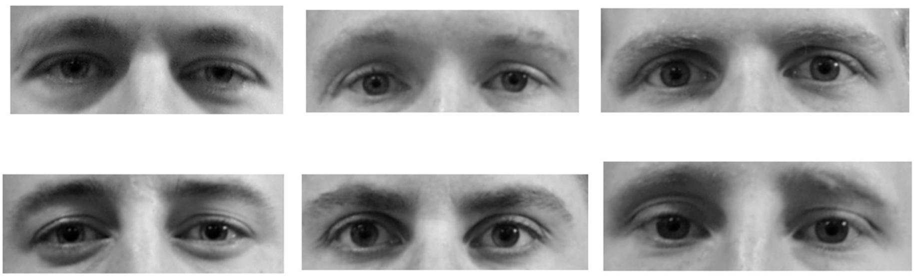

Figure 3. Distinctive face regions used in Experiment 2.

Procedure

The procedure was identical to that of Experiment 1.

Results and discussion

Table 2 shows the mean proportion of correctly identified stimuli in the test phase, again separately for those participants who learned the full faces and for those who learned the isolated features.

The data were submitted to a repeated measurement three-way ANOVA with learn size (learn–full vs. learn–part) as between-subjects factor and test size (test–full and test–part) as well as orientation (upright and inverted) as within-subjects factors. The analysis revealed a significant main effect of orientation, $F(1, 14) = 12.466$, $p < .01$, with upright being easier (upright, .672; inverted, .516). Moreover, there was an interaction between test size and learn size condition, $F(1, 14) = 25.671$, $p < .01$ (see Figure 2). No other effects were significant.

As in Experiment 1, whole-to-part superiority for upright faces was found only when full faces had been learned, one-tailed t test with $t(7) = 3.212$, $p < .01$, but again was reversed after learning isolated faces, as in Experiment 1 revealed by a two-tailed t test, $t(7) = 2.411$, $p < .05$; we used a two-tailed instead of a one-tailed t test due to a lack of a hypothesis of the direction of the effect.

This result supports the hypothesis that it is difficult to parse full faces into parts without being influenced by the irrelevant context. Even when the critical part is known, perceivers can hardly ignore the additional information. Thus, Experiment 2 again reveals that the full-face context does not provide an advantage in general. On the contrary, it may also hinder, even when the perceiver knows on which part of the face to focus. These interference effects in the part—whole conditions of Experiment 2 show that holistic interference effects cannot be avoided. Moreover, performance was even better when parts were learned (as compared to learn–full).

Together, Experiments 1 and 2 reveal two effects. Wholes are recognized better only when wholes are learned; they do not reveal superiority when isolated features are learned. Moreover, both experiments demonstrate that the participants were not able to ignore irrelevant features, when they had only learned parts. This seems to indicate that it is impossible to process the parts independently from the whole. The context seems to distract the perception of the parts, independently of the participants' ability to focus on a particular facial region.

We used stimuli in Experiments 1 and 2 in which additional features were well controlled in their context of "extra" information. These artificial faces allowed us to construct

TABLE 2
Mean identification rates of Experiment 2

	Test							
	Test–full–up		*Test–part–up*		*Test–full–in*		*Test–part–in*	
Learn	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Full	.812	.19	.625	.15	.521	.19	.417	.28
Part	.604	.30	.896	.15	.542	.25	.667	.20

faces that differed only in one critical part (such as eye, nose, or mouth) and were otherwise identical.

Real faces are less redundant than the artificial stimuli used in Experiments 1 and 2. In real faces all parts are more or less unique. Real faces differ not only from each other in respect to eyes, noses, and mouths as the stimuli in Experiment 1 and 2, they also differ in a number of other dimensions such as hairstyle, chin, cheekbones, texture, ears, and so on. Therefore, using real faces in Experiments 3 and 4 will reveal whether the effects generalize to stimuli with realistic variation.

EXPERIMENT 3

Pilot studies revealed that stimulus sets as small as those used above were too easily learned when using real faces, producing ceiling effects. Therefore, in Experiment 3, the number of stimuli was increased to nine, in order to make the task similar in difficulty to those in Experiments 1 and 2. Experiment 3 is very similar to Experiment 1 in that the critical parts in the isolated learning condition varied between faces.

Method

Participants

A total of 16 graduate students and undergraduates (13 females, overall mean age 27.1 years) from the Freie Universität Berlin participated for course credit. All participants were tested individually.

Materials

Nine real greyscale photographs were used. All were taken from male adults and are shown in Figure 4. Common German names (as in Experiments 1 and 2) were randomly assigned to the face stimuli. The names were Tobias, Sascha, Robert, Reiner, Konrad, Justus, Klaus, Peter, and Franz for the first name set. The second name set comprised Alfred, Arnold, Dieter, Gregor, Hannes, Jürgen, Simon, Anton, and Lukas.

Procedure

Participants had to learn three faces by their eye regions, three by their nose region, and three by their mouth region (nine parts in the learn–part learning condition), or they learned full faces (learn–full). In the test phase, all full faces and the specific parts were presented twice, either in an upright or in an upside-down position, thus yielding 72 trials per participant.

Results and discussion

Table 3 shows the mean proportion of correctly identified stimuli in the test phase, separately for the two learning versions, both test versions, part and whole, as well as both orientations at test.

A three-way repeated measurement ANOVA with learn size (learn–full vs. learn–part) as a between-subjects factor and test size (test–full vs. test–part) and orientation (upright vs.

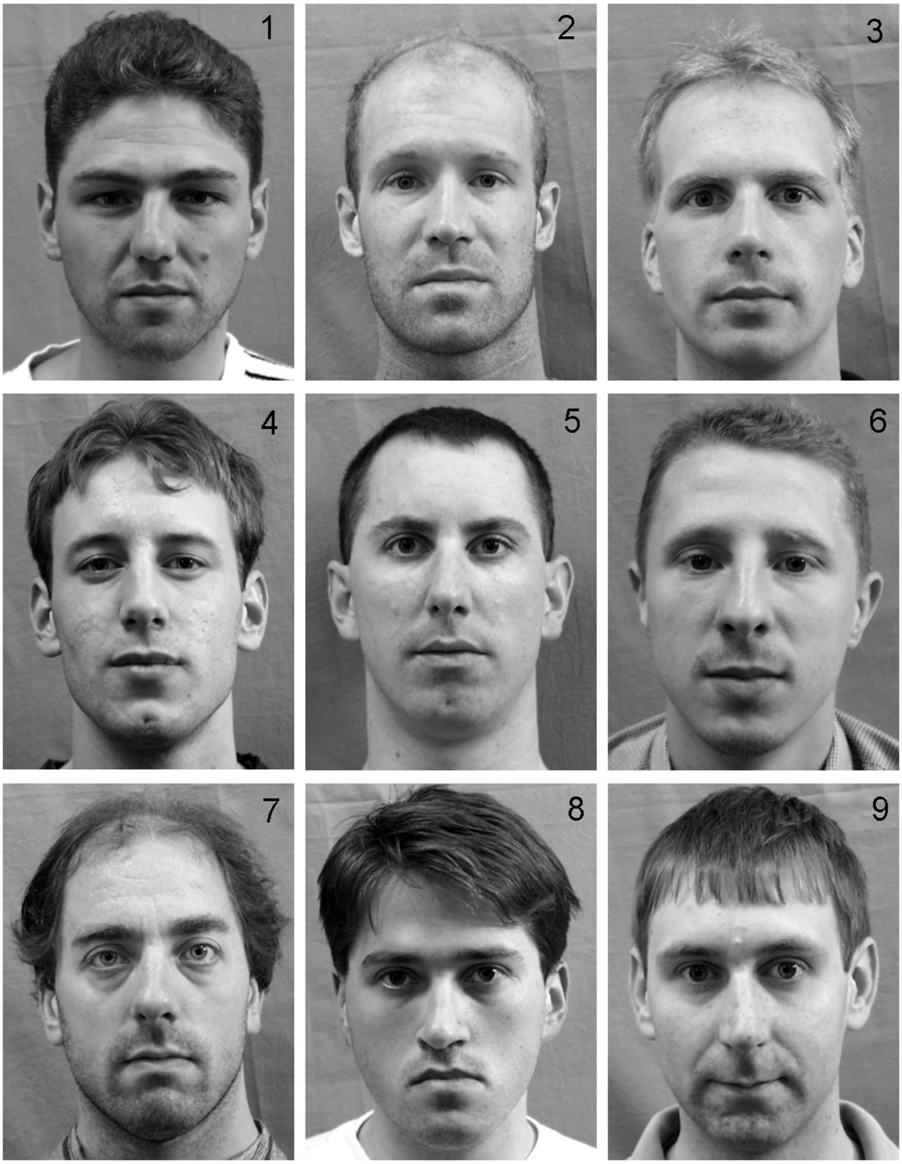

Figure 4. Face stimuli used in Experiment 3 and Experiment 4.

inverted) revealed a trend for a main effect of orientation, $F(1, 14) = 3.687$, $p = .0754$, *ns*, a significant effect of test size, $F(1, 14) = 18.075$, $p < .01$, and a significant interaction between test size and learn size, $F(1, 14) = 97.090$, $p < .01$ (see Figure 2). No other effects were significant.

Further analyses revealed that both simple main effects of test size under the two learn size conditions were significant. The effects were parallel to those already described in

TABLE 3
Mean identification rates of Experiment 3

	Test							
	Test–full–up		*Test–part–up*		*Test–full–in*		*Test–part–in*	
Learn	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Full	.833	.12	.354	.14	.820	.18	.319	.10
Part	.570	.23	.792	.19	.521	.17	.687	.11

Experiments 1 and 2: There was a dramatic decrease in performance for the learn–full condition when the participants were tested in the isolated test size condition. While their recognition for full faces in the full learning condition (.826) was very good, subjects' proportion of correct recognition for part test versions was low (.337). This difference was found to be significant by a one-tailed t test with $t(7) = 9.957$, $p < .01$.

However, most important, when parts had been learned there was again interference from the whole face at test, two-tailed t test with $t(7) = 3.968$, $p < .01$. Thus the real faces used here also showed the holistic interference effects found in Experiments 1 and 2. In order to test the effect of focusing to one specific facial region or feature as in Experiment 2 we tested the eye regions only in Experiment 4.

EXPERIMENT 4

Experiment 4 was very similar to Experiment 2 in that nine different eye regions were learned and tested in the isolated learn condition, and whole faces in the full condition. The same nonredundant real faces as those in Experiment 3 were used.

Method

Participants

A total of 16 graduate students and undergraduates (all females, overall mean age 26.8 years) from the Freie Universität Berlin participated for course credit or were paid for their participation. All participants were tested individually.

Materials

The same nine real faces and names as those in Experiment 3 were used. The procedure was similar to those in Experiment 1, except that the learn size condition part consisted only of isolated eye regions.

Results and discussion

The mean proportion of correctly identified stimuli in the test phase, separately for the two learn size versions and for both orientation and both test size conditions, sampled over all participants, is shown in Table 4.

Data were submitted to a three-way repeated measurement ANOVA with learn size

TABLE 4
Mean identification rates of Experiment 4

	Test							
	Test–full–up		*Test–part–up*		*Test–full–in*		*Test–part–in*	
Learn	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Full	.944	.05	.556	.16	.937	.10	.486	.17
Part	.701	.16	.771	.13	.639	.16	.667	.16

(learn–full vs. learn–part) as a between-subjects factor and test size (test–full and test–part) and orientation (upright and inverted) as within-subjects factors. A significant main effect of orientation at test, $F(1, 14) = 6.457$, $p < .05$, and a main effect of test size, with $F(1, 14) = 64.840$, $p < .01$, was found, as well as an interaction between learn size and test size condition, $F(1, 14) = 103.363$, $p < .01$ (see Figure 2).

In order to further investigate the interaction between learn size and test size, the whole-to-part superiority for learn–full faces as well as the part-to-whole superiority for learn–part faces, was analysed. When full faces had been learned the recognition rates for full faces at test were better than those for part-based versions at test. A one-tailed t test revealed that there was a whole-to-part superiority at test under the learn–full condition, $t(7) = 10.473$, $p < .01$. Nevertheless, after learning isolated parts, participants were slightly better in recognizing the facial parts without the context (see Table 4), two-tailed t test with $t(7) = 2.143$, $p = .0693$, *ns.*

Thus, perhaps because of the lower level of redundancy, this effect was smaller in Experiment 4 than it was in Experiment 2. We tested this with an additional three-way ANOVA with experiment (Experiment 2 vs. Experiment 4) as a between-subjects factor and test size (test–full vs. test–part) and orientation (upright vs. inverted) using the data from the part-learn conditions only. The analysis revealed an interaction of experiment and test size, $F(1, 14) = 7.310$, $p < .05$, which shows that a lower level of redundancy in real faces allows subjects more easily to ignore irrelevant information (see Figure 5). However, even with real faces dependencies of the learning condition were found.

Compared with Experiment 3 (see Figure 2), it seems that participants in Experiment 4 were better able to ignore the irrelevant context. To test this, we ran an additional repeated measurement three-way ANOVA with the between-subjects factor experiment (Experiment 3 vs. Experiment 4), test size (test–full vs. test–part) and orientation (upright vs. inverted) as within-subjects factors. Only the data from the part-learn conditions were used. If participants were better able to ignore the irrelevant context in Experiment 4, then there should be an interaction between experiment and test size. This interaction was found, $F(1, 14) = 7.272$, $p < .05$, and is illustrated in Figure 5. The advantage for the part-based faces against the full faces at test (in the part-based learn condition) was obviously smaller in Experiment 4 than in Experiment 3: Experiment 4, .049; Experiment 3, .194; two-tailed $t(14) = 2.697$, $p < .05$. This difference reveals the amount of interference caused by the context when focusing is made more difficult.

The findings are in accordance with a theory, which explains not only inversion effects, but also the whole–part results in Experiments 3 and 4. Real faces, as long as they are rather

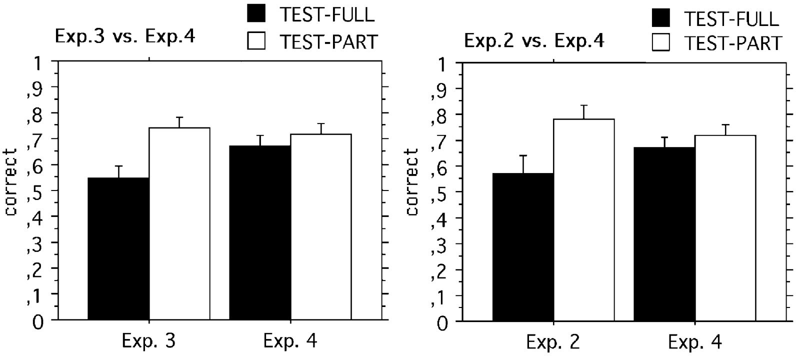

Figure 5. Proportion of correct rates after learning parts for Experiment 3 compared with Experiment 4 (left chart) and for Experiment 2 compared with Experiment 4 (right chart). Data are sampled over upright and upside-down conditions.

unfamiliar, are predominantly processed using the gross external features (Young et al., 1985). If participants in the learn–full conditions used this strategy, then the better performance in the full–full than in the full–part conditions is probably caused by a deficit of learning the individual parts in favour of hairstyle and distinctive head shapes, which are only present in the learn phase.

Concerning holistic interference effects the results of Experiment 4 reveal a very interesting though somehow unexpected effect. It was easier to ignore the irrelevant wholes at test when real faces were used (as indicated by the difference between part and full). When part-based versions had been learned, the part-to-whole superiority was .208 for Experiment 2, but only .049 for Experiment 4, two-tailed t test, $t(14) = 2.704$, $p < .05$. This is illustrated in Figure 5. It seems that the use of shared features in nearly all previous experiments testing holistic processing (Farah et al., 1995; Leder & Bruce, 2000; Tanaka & Farah, 1993) reduces the discriminatory ability of the cognitive system. Thus, Experiment 4 also reveals that interference seems to be smaller when the irrelevant parts are not redundant and are not identical over different trials.

GENERAL DISCUSSION

The present experiments shed new light on the sensitivity of the effects of the whole-to-part relationship in face recognition. In four experiments, participants memorized either full faces or facial parts. We replicated whole-to-part superiority effects. More important, after learning parts we showed how difficult it is to ignore irrelevant parts. We believe that this kind of interference might be the essence of holistic face processing.

In all four experiments, when full faces were learned, recognizing full faces was superior, and parts were difficult to parse from the whole. This was particularly difficult when it was not explicitly stated which of the components of the whole was critical.

Participants performed better with full faces in the test when they had learned full faces, and they did better in the test with parts only when they had learned the parts. A strict interpretation of the holistic processing hypothesis would lead to the assumption that wholes are superior in general. Apparently, this is not in line with our results. The "superiority" of the whole emerges only when the whole corresponds to a representation in memory. The decline of performance when full faces had been learned and recognition of parts was tested is consistent with predictions of the holistic processing hypothesis in that it shows that partitioning a face into its parts is difficult.

Most interesting were the conditions in which parts had been learned. Faces can be partitioned into meaningful parts. In everyday life, people tend to answer "eyes, nose, mouth" when asked what the constituting elements in faces are. Even the direct tests for a holistic processing hypothesis revealed that isolated parts are recognizable (Tanaka & Farah, 1993), though to a smaller extent when they were embedded in facial wholes.

In our experiments, when facial parts had been learned, a pure part-based approach would predict (if it is assumed that the parts discriminate successfully between different objects) that adding additional features would not affect the recognition of the object. Nevertheless, our studies showed different results: Learning parts brought about very good performance with parts. However, when the parts were embedded in a full face, recognition decreased dramatically (at least in Experiments 1, 2, and 3 and to a smaller extent in Experiment 4). The context of the whole interferes with the recognition of its parts. This is strong evidence against a general whole-to-part superiority and is in accord with the pattern of results reported by Homa, Haver, and Schwartz (1976), who used simple schematic faces. However, we believe that our results demonstrate a kind of holistic interference, which has not been shown in the same way before. However, the interference shown here bears some similarity with the findings of Farah et al. (1998) and reveals that irrelevant information, is hard to ignore in a recognition paradigm. Murray and Jones (2002) argued that interference effects might occur because irrelevant information sometimes is automatically processed up to a semantic level.

Whether the superiority of the whole face was due to additional information, or whether it indicates a different quality, cannot be decided conclusively from our data. The general level of performance does not support the hypothesis that it is the whole that makes faces particularly well recognizable stimuli, as there are conditions in which parts were better recognizable (see results of Experiments 1, 2, and 3).

Another important finding of our study was that even when participants knew that the eye region was the critical feature (in the learn–part conditions of Experiments 2 and 4), there was still a decrease in performance with the test–full faces (as can be seen from Figure 2). Thus, it is the influence of the unexpected context that hinders recognition, not the uncertainty about the critical parts alone.

Concerning inversion effects, the experiments using real faces (Experiments 3 and 4) showed no specific inversion effects, and the effect of orientation was weaker for the real faces than in the first two experiments. However, the stimulus variation used here gives the real faces in Experiments 3 and 4 some advantage. The tasks of Experiments 3 and 4 were probably easier, and all face conditions contained information that was less affected by inversion than the more redundant—and, therefore, more difficult—information in Experiments 1 and 2. Real faces differ in respect to hairstyle, chin shape, and so on, which are all gross features that are possibly not affected as much by orientation as are internal features, (Sergent, 1984). However, the processing of artificial faces, which differ only in respect to internal features, bring out stronger effects of locally processed configural features. Leder and Bruce

(1998), for example, showed that effects of distinctiveness based on manipulations of local or configural internal facial features come out much stronger when the outer features had been omitted.

In our study concerning wholes and parts, inversion produced differential effects. However, interpretation of these effects is difficult, because the inversion effects can be expressed in terms of absolute or relative decrease. Experiment 1 showed that inversion deficits were particularly large in full faces at test when these have already been learned as full faces. This finding supports both a holistic and a configural explanation of the face inversion effect. The larger inversion effects in learn–full conditions support the hypothesis that inversion effects depend on the availability of configuration or holistic representations, which are available to a much greater extent from full faces than from parts. On the other hand, inversion deficits for parts were larger when parts had been learned. This cannot easily be explained by the dependence of inversion effects on configural holistic processing. Rather it supports a learn–test compatibility explanation (Tulving & Thomson, 1973), according to which the learned conditions are mainly superior, and inversion sometimes disrupts their efficient processing, be they parts or wholes. This is in accordance with the results of Experiment 2 in which there were no interactions with orientation when it was possible to focus on one facial region. In Experiments 3 and 4 the effects were more ambiguous because of the high performance in some conditions and the generally small inversion effects for full faces (especially in Experiment 3), which are probably due to the dominance of gross facial features, which might be less affected by inversion.

Another interesting finding concerns the hypothesis that eye regions contain more locally processed relations than do other face regions. Leder, Candrian, Huber, and Bruce (2001) found that eye distance is processed as a local relational feature and that its processing is disrupted by inversion. In accordance with this hypothesis, inversion effects for the part versions at test in Experiment 2 (when eye regions alone were used) are larger than those in Experiment 1: Experiment 1, .073; Experiment 2, .219; one-tailed $t(30) = 1.930$, $p < .05$. However, in Experiments 3 and 4 (see Table 4) there was only a marginal difference between these conditions, $t(30) < 1$, *ns*. Understanding these findings requires additional research in the future.

In Experiments 1 and 3, we did not analyse the recognizability of the three different face regions. We had no hypotheses concerning the different regions although studies about feature saliency in faces would predict that eye regions might be better recognized (Shepherd, Davies, & Ellis, 1981). In our experiments variation in the target features was used to test the focusing hypotheses. Our results revealed that holistic processing was stronger when subjects did not focus on one particular face region. Only in Experiment 4 did it seem as if it was possible to ignore the context relatively successfully.

Interference was stronger when the artificial faces in Experiments 1 and 2 were used. However, the findings with real faces in Experiments 3 and 4 also revealed the critical learn–test interactions. Interestingly, Experiment 4 revealed that in real faces it seems to be easier to ignore irrelevant parts, which certainly has implications for interpretations made from the experimental literature based on artificial faces.

By comparing results with real and artificial faces we were able to support our findings using different levels of resemblance to everyday face recognition. To summarize, our findings show that the whole face can certainly hinder the recognition of its parts. Interference of the whole face on the processing of its parts presumably is an essential feature in holistic processing. Part–whole interference effects therefore might be used as a possible indicator of holistic processing.

REFERENCES

Donnelly, N., & Davidoff, J. (1999). The mental representations of faces and houses: Issues concerning parts and wholes. *Visual Cognition, 6*, 319–343.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 628–634.

Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*(3), 482–498.

Homa, D., Haver, B., & Schwartz, T. (1976). Perceptibility of schematic face stimuli: Evidence for a perceptual Gestalt. *Memory and Cognition, 4*(2), 176–185.

Leder, H. (1998). *The ATR face databank*. Unpublished manuscript.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology, 51A*(3), 449–473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology, 53A*(2), 513–536.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*, 73–83.

Murray, J. E., & Jones, C. (2002). Attention to local form information can prevent access to semantic information. *Quarterly Journal of Experimental Psychology, 55A*(2), 609–625.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*, 221–242.

Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces*. London: Academic Press.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46A*(2), 225–245.

Tanaka, J. W., & Gauthier, I. (1997). Expertise in object and face recognition. In R. L. Goldstone, P. G. Schyns, & D. L. Medin (Eds.), *Psychology of learning and motivation, mechanisms of perceptual learning* (Vol. 36, pp. 83–125). San Diego, CA: Academic Press.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 359–380.

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception, 14*, 737–746.

Young, A. W., Hellawell, D., & Hay, D. (1987). Configural information in face perception. *Perception, 10*, 747–759.

Original manuscript received 14 March 2002
Accepted revision received 22 September 2003