Comparing inverted faces to upright faces using similarity or mental rotation

Sam S. Rakover, Rani Amit Bar-On, and Anna Gliklich

Department of Psychology, Haifa University, Haifa, Israel 31905

Running head: Inverted face vs. upright

Rakover phone number: 972 4 8240924

Email: [rakover@psy.haifa.ac.il](mailto:rakover@psy.haifa.ac.il)

Correspondence should be addressed to Sam S. Rakover, Department of Psychology, Haifa University, Haifa, Israel 31905.

Email: [rakover@psy.haifa.ac.il](mailto:rakover@psy.haifa.ac.il)

**Abstract**

A major interest of research in face recognition lies in explaining the Face Inversion Effect (FIE), in which the recognition of an inverted face is less successful than that of an upright face. However, prior research has devoted little effort to examining how the cognitive system handles comparison between inverted and upright faces. **The results of a preparatory experiment and two following experiments (experiment 2a and 2b) support the conclusion that the comparison is based more on visual similarity than on a** mental rotation of the inverted face to an upright face. Visual similarity is based on certain elements mutual to the two faces, which resist the transformation of inversion. These elements are symmetrical or salient components of the face, such as round eyes or thick lips.

Keywords: Face recognition, face inversion effect (FIE), visual similarity, mental rotation.

Comparing inverted faces to upright faces using similarity or mental rotation (Note 1)

Research on face perception and recognition over the last 50 years has fo­cused on the Face Inversion Effect (FIE), according to which an upright face (with the hair on top and the chin below) is recognized much better than an inverted face (chin on top, hair below, at **a rotation of 180 degrees**) ( **Maurer, Le Grand, & Mondloch, 2002**; Rakover, 2002, 2013; Rossion, 2008, 2009; Valentine, 1988; Yin, 1969).  **This research has concentrated on four types of facial information: featural (eyes, nose, and mouth), relational (eyes above nose and nose above the mouth), configural (space between eyes, as well as space between nose and mouth), in addition to a holistic perception of the face. Two hypotheses that arise from this research—the configural processing and holistic hypotheses—suggest that all four types of informa­tion undergo appropriate processing when a face is presented upright. However, when the face is presented upside down, configural and holistic information are greatly impaired, but featural information processing remains intact (Maurer et al., 2002; Rakover, 2002, 2013; Rossion, 2008, 2009; Valentine, 1988; Yin, 1969). Hence, while in­verted face processing is part-based (featural), upright face processing is principally configurally and holis­tically based (Maurer et al., 2002; McKone, 2010; Piepers & Robbins, 2012; Rakover, 2013; Rossion, 2008, 2009).**

**Nevertheless, there are two notable issues with the explanation of the FIE. Firstly, several researchers have found that configural information is not a crucial condition for FIE: the FIE can be obtained with isolated features (e.g., Rakover, 2012; Rakover and Teucher, 1997), inversion does not impair previous changes in the configural information that generated the eye-size illusion (e.g., Rakover, 2011; see however Xiao et al., 2014 for critique and Rakover, 2017 for response), and empirical findings suggest that featural and configural information affect recognition of inverted faces (e.g., Civile, McLaren & McLaren 2016).**

**Another issue is that previous research on the FIE did not answer the question of how the cognitive system compares a perceived inverted face with that of a remembered upright face. The present paper attempts to answer this question. To locate a possible answer, we tested two hypotheses: visual similarity and mental rotation.**

The **visual-similarity** hypothesis proposes that a person’s decision is based on similarity of mutual alike elements between the perceived inverted face and the remembered upright face (Rakover & Cahlon, 1989, 2001; Tversky, 1977). The **mental-rotation** hypothesis proposes that the inverted face as a whole is mentally rotated to the upright orientation and then compared to the remembered upright face (Rakover, 2015; Rock, 1973, 1974; Valentine & Bruce, 1988). While the FIE can be explained by the great difficulty of mentally rotating each of the facial features to the upright orientation (Rock, 1973, 1974), empirical evidence supports the hypothesis that a face is mentally rotated as a whole unit, similar to the mental rotation of entire visual shapes (Cooper, 1975; Shepard & Metzler, 1971; Valentine & Bruce, 1988).

These two hypotheses do not overlap since they stem from two distinct mechanisms. While the visual-similarity hypothesis is based mainly on the estimation of the number of mutual elements that compound the two faces (Rakover & Cahlon, 1989, 2001; Tversky, 1977), the mental-rotation hypothesis is founded on a mechanism that rotates the representation of the inverted face to the upright orientation and then examines whether it overlaps with the remembered upright face (Cooper, 1975; Rakover, 2015; Rock, 1973, 1974; Shepard & Metzler, 1971; Valentine & Bruce, 1988). Several studies suggest that object recognition relies on similarity, whereas mental-rotation is based on the angular disparity between the two tested stimuli (objects). Furthermore, these two mechanisms are associated with two different areas of the brain (Cheung et al., 2009).

The main goal of the present paper is to empirically test these two hypotheses (for a consideration of a combination of these hypotheses, see the Discussion section). We conducted two experiments based on the following manipulations and rationale. First, in a preparatory experiment, we constructed two groups of faces. The *similar* group contained seven different pairs, each composed of two different faces that had obtained a high similarity index, one upright and one inverted. The *non-similar* group contained seven different pairs each composed of two different faces, one upright and one inverted, which had obtained a low similarity index.

Secondly, in a prediction testing experiment, a variation of the Yes/No recognition procedure, we presented to the participants in a *Study* stage, 14 upright faces taken from both the *similar* and *non-similar* groups. In a subsequent *Testing* stage, 28 inverted faces were presented. They were composed of the 14 previously viewed upright faces (of the *Study* stage) and 14 new faces. The 14 new faces included seven inverted faces from the *similar* group and seven from the *non-similar* group. The participant’s task was to decide for each inverted face if it was old or new.

If the **visual-similarity** hypothesis is correct, we can assert the following specific prediction: The percentage of false-alarms for the seven new similar inverted faces (%FAs) will be significantly greater than the percentage of false-alarms for the seven new non-similar inverted faces (%FAns). When the visual similarity between upright face and inverted face is high, there is a greater likelihood of believing that a new face is an old one; hence, %FAs > %FAns.

In contrast, if the **mental-rotation** hypothesis is correct, there will be no significant differences between %FAs and %FAns; when an inverted face is rotated to the upright position, it should be easy to decide whether or not it is congruent with the remembered upright face and therefore to decide whether the inverted face is old or new. Moreover, given that (a) reaction time and the number of errors increase as a function of the angular disparity between the two stimuli (Cheung et al., 2009; Cooper, 1975; Shepard & Metzler, 1971), and that (b) the angular disparity was held constant in the present experiment (inverted faces are rotated to the upright position), we make the following specific prediction: There will be no significant differences between %FAs and %FAns since the angular disparity is the same.

To test these two predictions, we conducted two experiments. The goal of the preparatory Experiment 1 was to construct two groups of faces: *similar* and *non-similar*. The testing Experiment 2 had two parts. The first, *Similarity,* was to test the above two predictions empirically, and the second, *Orientation,*had two important goals. First, to give additional empirical support to the construction of the *similarity* groups by using the technique of similarity ranking, and second, to test whether the distinction between the *similar* and the *non-similar* groups was confined only to the pair of faces in the Upright/Inverted (UI) orientations or if it could be generalized to the other orientations, UU, IU and II. In other words, Part B of Experiment 2 was to test whether this distinction has the quality of resisting the inversion transformations.

Experiment 1 – Preparatory study

Methods

The study has been reviewed and approved by the Psychology department ethics committee and the University of Haifa’s IRB (application no. 234/17). All methods were carried out in accordance with relevant guidelines and regulations. The materials and data are available at <https://drive.google.com/drive/folders/1Vxo48qsm742JVGE1H_TWBKV0K9A7Lt7s?usp=sharing>.

The black and white face-pictures used in the present study were taken from 2 publicly available online data banks back in the 1990’s. **A selection of these pictures was utilized in several previous experiments (e.g., Rakover, 2015, 2017).**

*Participants, Design, and Procedure*: Thirty participants (21 females and nine males, average age 24.7) were shown a series of face pictures on a computer screen **(AOC company, G2460PF model, 24 inches, X10801920 resolution)**. The number of participants was determined on the basis of a pilot study and previous experience with similar kinds of experiments. The participants were undergraduate students who were rewarded by payment or course credit. Informed consent was obtained for the experiments reported here. In each trial, participants were presented with six oval-shaped faces without hair or ears. One of the faces appeared in the upright orientation (hair on top, chin below) and the other five appeared in the inverted orientation (chin on top, hair below) and were arranged in a semi-circle below the upright face. Each oval face was a different unfamiliar, black-and-white image of a man. The hair and ears were cut because external facial features would aid recognition (Bonner et al., 2003; Want et al., 2003).

A total of 180 trials included 30 different faces appearing six times each in the upright orientation with five different inverted faces below. Thus, each upright face was associated with all 30 faces in the inverted orientation (including itself). The faces, chosen randomly for each trial, were exposed for 20 seconds, at which time the participant had to choose among the five inverted faces at least one similar to the upright face. Each trial also featured two whistles, at 10 and 18 seconds, to expedite the participant’s responses.

Results

The main purpose of the preparatory Experiment 1 was to construct two groups of face pairs: *similar*and *non-similar*. To do this, we constructed a table of 30 *upright* faces x 30 *inverted* faces. The table has 30x30=900 cells. Each cell shows how many of the 30 participants indicated that a given inverted face was similar to a certain upright face. For example, cell4,15 = 15 means that half of the participants indicated that inverted face4 is similar to upright face15. Based on this result, we built a *similar*group by selecting seven different pairs of different faces (upright and inverted) with a high similarity index (similarity was in the range between 27% and 67% of the participants). The *non-similar* group was constructed in the same way; we selected seven different pairs of different faces (upright and inverted) with a low similarity index (similarity was in the range between 3% and 17% of the participants). The pairs in these two groups were different from each other. Figure 1 shows examples of the *similar* and *non-similar* groups.

=========================

Insert Figure 1 about here

===========================

It was found that on average, 82% of participants indicated that an inverted face was similar to itself in the upright orientation. In comparison, an average of 23% of participants indicated that a certain inverted face was similar to a differentupright face. With regard to similarity, this finding suggests that the number of elements mutual to upright face X and inverted face X is greater than the number of elements mutual to upright face X and inverted face Y.

Experiment 2 – Prediction testing

Methods

Part A – Similarity

*Participants, Design, and Procedure*: Twenty new participants (15 females and five males; average age = 24.75) were recruited and were shown a series of pictures on a computer screen; each picture consisted of one face. The experiment consisted of two stages: *Study* and *Test* and was a variation of the common Yes/No recognition experiment. In the *Study* stage, participants were shown 14 upright faces, each exposed for 3s at intervals of 1s between faces. These upright faces were taken from both the *similar* and *non-similar* groups and presented in randomized order. In the *Test* stage, participants were shown 28 inverted faces including 14 *old* faces that appeared in the *Study* stage and 14 *new* faces taken from both the *similar* and *non-similar* groups. The faces in this stage were also presented in randomized order.

The participants were informed that they would take part in two experiments in succession. The instructions for each experiment were read before the beginning of each experiment. For Part A, the *Study* stage, participants were told to concentrate on the upright oval faces and attempt to remember them. In the *Test* stage, they were presented with *old* and *new* inverted faces, one at a time, and had to decide whether they were *old* or *new*. They had 10s to make their decision. After 5s, a whistle sounded to expedite their decision. **Based on previous experiments, a lengthy decision time was utilized to relieve the participants from time pressure.**

Part B – Orientation

*Participants, Design & Procedure*: The same 20 participants were shown a series of pairs of faces, one at a time, on a computer screen. All the pairs were constructed of the14 pairs constructed in the preparatory study, Experiment 1. Each face pair was presented in four possible orientations: *upright/inverted* (UI), *upright/upright* (UU), *inverted/inverted* (II) and *inverted/upright* (IU). A total of 56 face pairs (4x14) were presented one at a time in randomized order. Each pair was presented for 10s to allow the participant to rank the degree of similarity between the faces on the screen; a score of 1 signifies that the faces were not similar at all, while a 5 signifies that the faces were very similar. After 5s, a whistle sounded to indicate that half the time remained.

Results

*Part A* The main results appear in Figure 2. While there is no significant difference between Percent Hits in the *similar* groups (%Hs) and Percent Hits in the *non-similar* groups (%Hns), Percent False-Alarm in the similar-groups (%FAs) is significantly greater than Percent False-Alarm in the non-similar-groups (%FAns). A repeated measurement 2 (Hits, False-Alarm) x 2 (*similar* group, *non-similar* group) ANOVA supports this observation: F(1,19) = 9.56 p<.006 µ2=.34; A LSD test revealed a significant difference between %FAs=55.0% and %FAns=35.7% p<.007 but not between %Hs and %Hns.

=============================

Insert Figure 2 about here

=============================

*Part B The* main results appear in Figure 3. The similarity ranking of the pairs in the *similar* group (SRs) was higher than the similarity ranking in the non-*similar* group (SRns) in all the four groups: UI, UU, IU and II. A repeated measurement 4(UU, UI, IU, II) x 2(Similar group, Non-similar group) ANOVA supports this observation: F(3, 57) = 27.92 p< .0001 µ2=.595. An LSD test done within each of the four groups revealed that SRs was significantly higher than SRns at the level of p<.001.

=============================

Insert Figure 3 about here

=============================

Discussion

The main results of the present study are as follows: First, %FAs is greater than %FAns, while no significant difference was found between %Hs and %Hns. Second, the SRs was higher than the SRns in each of the four orientations. The first result tends to support the visual similarity hypothesis, such that the cognitive systemcompares an inverted face with an upright face by conducting a visual similarity between the perceived inverted face and the remembered upright face (Rakover & Cahlon, 1989; Tversky, 1977). The result does not support the **mental-rotation** hypothesis since it predicts no significant difference between %FAs and %FAns.

Given this, one may propose the “combined hypothesis”, a kind of a combination of mental rotation and visual similarity; the inverted face is rotated to the upright orientation and then visual similarity is activated. One important argument against this hypothesis is that it involves a huge mental effort; one has to rotate the perceived face to the upright orientation and then compare it with the remembered upright face to evaluate their visual similarity. Conversely, according to the visual-similarity hypothesis, one can simply evaluate the visual similarity of the two faces (inverted and upright) directly. As can be seen from the results of Experiment 2, Part B, visual similarity resists the transformation of inversion. As an example, consider the II pair (see Figure 3). According to the combined hypothesis, the similarity evaluation of the II pair should be done indirectly, after mental rotation of each of the inverted faces to the upright orientation – an action that demands a great mental effort. In contrast, according to the visual-similarity hypothesis, the evaluation is direct and does not demand so much mental effort. Since exerting high levels of mental effort tends to produce errors, one may predict that the similarity evaluation of the II pair will be the lowest; however, as can be seen in Figure 3, the similarity evaluation is highest for this pair [*F*(3,57) = 29.17 *p* < .001 µ2 = .61]. Hence, one may cast doubt on the combined hypothesis.

**The second result, the “generalized similarity-ranking”, supports the validity of the construction of the *similar*** and *non-similar* groups; the SRs is greater than the SRns for the UI orientation and also in the other three orientations. Hence, the distinction between the *similar* and *non-similar* groups is not confined only to the UI orientation. Rather, the distinction generalizes to all other orientations. The **generalized similarity-ranking** finding suggests that the similarity between different faces resists the transformation of inversion. If we assume that similarity is founded on mutual elements that resist the transformation of inversion, we may propose that these elements have the qualities of symmetry or saliency, for example, round eyes, fleshy lips, wide nose, and thick eyebrows. This explains well the above finding; the similarity of an upright, oval face with itself in the inverted orientation is much higher than the similarity between an upright face with a different inverted face. The number of mutual symmetrical or salient elements in an upright face X and an inverted face X is much higher than in an upright face X and an inverted face Y. Furthermore, these elements may also explain why recognition of inverted faces is based mainly on featural information (Rakover, 2002, 2013; Rakover & Cahlon, 2001).

**Given the generalized similarity-ranking, one may hypothesize that %FAs will be greater than %FAns not only in the UI orientation group but also in the UU, IU, and II groups. We began testing this hypothesis by running the appropriate experiments. However, as a result of the COVID-19 pandemic, we only succeeded in completing the UU experiment. This experiment is identical to Experiment 2: Part A – Similarity, except that in the *Test stage* all the faces were presented in the upright orientation (there were 14 females and 11 males; average age = 23.68.) The hypothesis has been confirmed since the (Hit, FA) x (similar, non-similar) interaction was significant [F91,24)=18.45, p<.001, µ2 = .44]: %Hns (77.7%) was greater than %Hs (65.1%), LSD p=.008, but %FAs (38.9%) was greater than %FAns (23.4), LSD p=.002. Here, we obtained a pattern of results that was similar to the pattern of results from the previous experiment 2: Part A.**

**This experiment suggests two notable conclusions. Firstly, since in the present UU experiment, mental rotation is not needed (and therefore a difference between %FAs and %FAns is not predicted), the present finding provides strong empirical support in favor of the visual-similarity hypothesis. Secondly, the findings that one can explain the results of experiment 2: Part A and of the UU-experiment by the same hypothesis, the visual-similarity, may propose that the explanatory mechanism involved in the FIE is visual-similarity: one first detected in the memory of the appropriate group of faces (those previously seen) and then s/he operated the visual-similarity mechanism, based on which a decision (the presented face is old/new) is reached. This mechanism generates errors as a result of when the similarity among faces increases, the confusion among them also increases (e.g., Alexander & Zelinsky, 2012). Hence, for the similarity group, a face previously seen, may increase mistakes, and a new face may increase false alarms.**

Finally, the following concern must be considered. Despite that the present study’s findings tend to support the visual similarity hypothesis and not the mental-rotation hypothesis, one may propose that this conclusion is confined to the particular research task. Whereas the process of mental rotation is based on an identity-judgment task (measured by reaction time), visual similarity is based on a yes/no recognition procedure (measured by the percentage of false alarms). A number of studies support the idea that mental rotation is used in identity-judgment experiments, whereas similarity assessments are used in experiments of recognition (Cheung et al., 2009; Cooper, 1975; Shepard & Metzler, 1971; Valentine & Bruce, 1988). Further experiments are needed to gain additional insight into this matter.**References**

Alexander, R. G. & Zelinsky, G. J. (2012). Effect of part-based similarity on visual search: The Frankenbear experiment. *Vision Research,*54, 20-30.

Bonner, L., Burton, A. M., & Bruce, V. (2003). Getting to know you: How we learn new faces. *Visual Cognition*, *10*(5), 527–536. https://doi.org/10.1080/13506280244000168

Cheung, O. S., Hayward, W. G., & Gauthier, I. (2009). Dissociating the effects of angular disparity and image similarity in mental rotation and object recognition. In *Cognition* (Vol. 113, Issue 1, pp. 128–133). Elsevier Science. https://doi.org/10.1016/j.cognition.2009.07.008

Civile, C., McLaren, R. & McLaren, Ian P.L. (2016). The face inversion effect: Role of first-and second-order configural information. *American Journal of Psychology*, 129, 23-35.

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, *7*(1), 20–43. https://doi.org/https://doi.org/10.1016/0010-0285(75)90003-1

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. https://doi.org/10.1016/S1364-6613(02)01903-4

McKone, E. (2010). Face and object recognition: How do they differ? In *Tutorials in visual cognition.* (pp. 261–303). Psychology Press.

Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms “holistic,” “configural,” and “relational” in the face perception literature. In *Frontiers in Psychology* (Vol. 3, Issue DEC). https://doi.org/10.3389/fpsyg.2012.00559

Rakover, S. S. (2002). Featural vs. Configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology*, *93*(1), 1–30. https://doi.org/10.1348/000712602162427

Rakover, S. S. (2011). Configural processing hypothesis and face-inversion effect. In Y. H. Zhang (Ed.), Advances in face image analysis: Technique and technologies (pp. 316–333). Hershey, PA: IGI Global.

Rakover, S. S. (2013). Explaining the face-inversion effect: The face-scheme incompatibility (FSI) model. *Psychonomic Bulletin and Review*, *20*(4), 665–692. https://doi.org/10.3758/s13423-013-0388-1

Rakover, S. S. (2015). Cognitive processing of scrambled faces: Effects of instructions and task. *American Journal of Psychology*, *128*(3), 379–386. https://doi.org/10.5406/amerjpsyc.128.3.0379

Rakover, S. S. (2017). The eye-size illusion is handled by two separate processes. *American Journal of Psychology*, 130, 329–337.

Rakover, S. S., & Cahlon, B. (1989). To catch a thief with a recognition test: The model and some empirical results. *Cognitive Psychology*, *21*(4), 423–468. https://doi.org/10.1016/0010-0285(89)90015-7

Rakover, S. S., & Cahlon, B. (2001). *Face Recognition : Cognitive and Computational Processes* (Vol. 31, Issue v. 31). John Benjamins Publishing Co. https://doi.org/10.1075/aicr.31

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole

relationship. *Perception & Psychophysics*, 59, 752–761.

Rock, I. (1973). *Orientation and form*. Academic Press.

Rock, I. (1974). The Perception of Disoriented Figures. *Scientific American*, *230*(1), 78–86. http://www.jstor.org/stable/24949985

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, *128*(2), 274–289. https://doi.org/10.1016/j.actpsy.2008.02.003

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: The perceptual field hypothesis. *Acta Psychologica*, *132*(3), 300–312. https://doi.org/10.1016/j.actpsy.2009.08.002

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science (New York, N.Y.)*, *171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327–352. https://doi.org/10.1037/0033-295X.84.4.327

Valentine, T. (1988). Upside‐down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*(4), 471–491. https://doi.org/10.1111/j.2044-8295.1988.tb02747.x

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory & Cognition*, *16*(6), 556–566. https://doi.org/10.3758/BF03197057

Want, S. C., Pascalis, O., Coleman, M., & Blades, M. (2003). Recognizing people from the inner or outer parts of their faces: Developmental data concerning “unfamiliar” faces. *British Journal of Developmental Psychology*, *21*(1), 125–135. https://doi.org/10.1348/026151003321164663

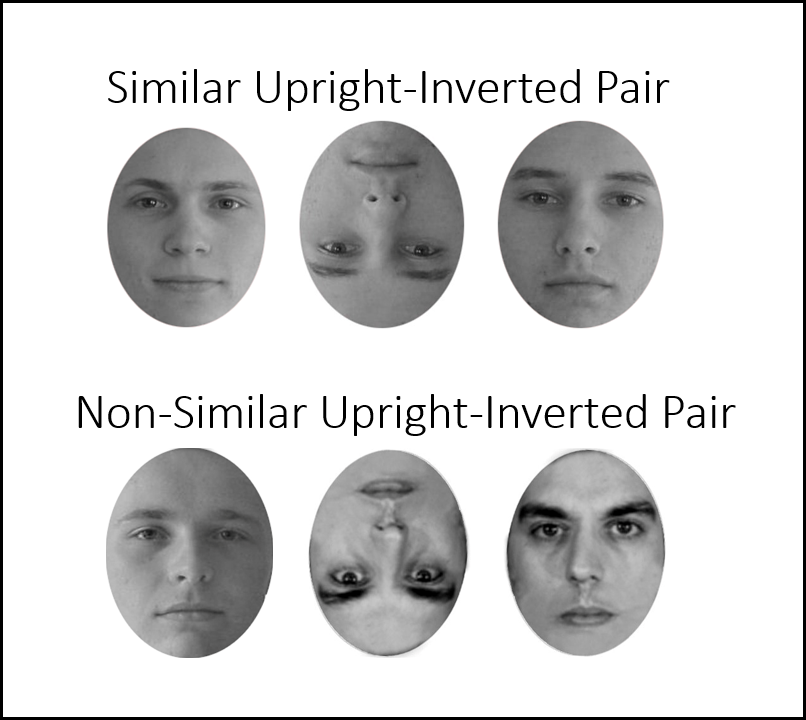
Xiao, W. S., Fu, G., Quinn, P. C., Sun, Y. H., Xiao, N. G., Wang, Q., . . . Lee, K. (2014). The eye-size illusion: Psy­chophysical characteristics, generality, and relation to holistic face processing. *Perception*, 43, 265–

274.

Yin, R. K. (1969). Looking at upide-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145. https://doi.org/10.1037/h0027474

**Notes**

1. **We like to thank the two anonymous reviewers and the journal editor Robert W. Proctor, for their helpful comments, which greatly improved the paper.**



**Figure 1:** Examples of similar and non-similar pairs of oval-faces. The left face was presented in the upright orientation and the middle one in the inverted orientation. The upright face on the right is the same as the inverted one. It is presented here for the sake of comparison.

**Figure 2:** PercentHits and False Alarm as function of similar and non-similar groups of pairs of different oval-faces.

**Figure 3:** Similarity ratings as a function of the face pair orientation and of similarity versus non-similarity.