Perception, 1999, volume 28, pages ***–***

From rotation to disfiguration: Testing a dual-strategy model for recognition of faces across view angles

Dominique Valentin*†, Hervé Abdi*, and Betty Edelman*

*** School of Human Development, The University of Texas at Dallas, Richardson, TX 75083-0688, USA.**
† Université de Bourgogne, 21004 Dijon Cedex, France.

Abstract. This study investigates the effect of distinctive marks on the recognition of unfamiliar faces across view angles. Subjects were asked to memorize a set of target faces, half of which had distinctive marks. Recognition was assessed by presenting the target faces, either in the same orientation, or after 90 degrees rotation, mixed with an equal number of distractors. Results show that the effect of distinctive marks depends on the view presented during learning. When a frontal view was learned, as predicted by the Valentin, Abdi, Edelman, and Posamentier (in press) dual-strategy model, distinctive marks improve recognition performance in the 90-degree condition but not in the 0-degree condition. However, when a profile view was learned, distinctive marks have no effect on recognition performance, even in the 90-degree condition where a frontal view is tested.

1. Introduction

How do human observers transfer information from a single view of a face to new views of this face? Experiments investigating this issue are generally called "transfer studies." They were performed, at first, in the late seventies to evaluate the ability of eyewitnesses to recognize a suspect from a different perspective than that of the first encounter (Patterson & Baddeley, 1977). Most of the transfer studies use a yes/no recognition paradigm in which the orientation and/or expression of faces is changed between learning and testing. During the learning phase, subjects are asked to memorize a set of target faces presented in a given orientation/expression. During the testing phase, they are presented with new views of the target faces either in the learned orientation/expression, or in a different orientation/expression mixed with an equal number of distractor faces. The subjects' task is to indicate which faces they have seen during the learning phase. Typically, signal detection methodology is used to evaluate the ability of subjects to discriminate between targets and distractors.

The results of such transfer tasks indicate that the ability to generalize from single views of faces depends on subjects' familiarity with the faces. If the subjects are *familiar* with a face, they have no problems recognizing it, even after a 90-degree rotation (Bruce, 1982; Valentin, Abdi, & Edelman, 1997). However, if subjects are *not familiar* with a face, their recognition performance declines as the face is rotated in depth. The magnitude of the decline of performance varies with the studies.

Early studies showed that recognition accuracy is *not* affected by modifications of the appearance of the faces between learning and test such as changes in head orientation from a frontal to a 3/4 view and *vice versa* and/or changes in expression from smiling to neutral (Davies, Ellis, & Shepherd, 1978; Patterson & Baddeley, 1977). In contrast, more recent studies found that recognition performance for unfamiliar faces is significantly impaired when faces are rotated by 45 degrees between learning and test (Baddeley & Woodhead, 1981; Bruce, 1982; Hill, Schyns, & Akamatsu, 1997; Krouse, 1981; Logie, Baddeley, & Woodhead, 1987; O'Toole, Bülthoff, Troje & Vetter, 1995; Valentin *et al.,* 1997). Moreover, the decline in performance observed for unfamiliar faces is not a linear function of the rotation between learning and test. For example, Patterson and Baddeley (1977) found that subjects identified accurately 90% of the faces when no rotation occurred between learning and test, 88% after a 45-degree rotation, but only 79% after a 90-degree rotation. Likewise, Valentin (1996) found that human observers' recognition performance remains relatively stable for changes in orientation from 0 to 30 degrees. After 30 degrees, performance drops off and reaches an asymptotic value of $d' = 1$.

In summary, transfer studies show that human observers are able to recognize unfamiliar faces across view angles with a level of accuracy better than chance. The problem, however, is that these studies do not provide information on *how* the subjects perform the task. Do they use the visual information contained in a given 2D view to infer what a face is likely to look like from another view point? Or, alternatively, do they extract, from the face, some informative characteristic visible from most viewpoints and use this information to perform the recognition task. To address this issue, Valentin, Abdi, Edelman, and Posamentier (in press) compared the performance of an autossociative memory (*i.e.,* a neural network model; see, *e.g.,* Abdi, 1988; Abdi, Valentin, & Edelman, 1999) and human subjects. Human subjects and the model tried to recognize faces seen with a different view angle at test than the view learned. Results (*cf.* Figure 1) show that the performance of both human observers and the autoassociative memory decrease significantly after a 30-degree rotation in depth. After this point, the autoassociative memory is at chance level ($d' = 0$) whereas human observers remain above chance ($d' = 1$).

Based on these results, Valentin *et al.,* (1998) proposed that human observers used two different strategies to perform a recognition task across changes in orientation. Up to 30 degrees, the similarity between original and rotated views would be large enough to allow a global matching strategy or "transfer by configuration." Beyond 30 degrees, the similarity between original and rotated views being too small for a global matching strategy to succeed, subjects would rely on very localized information visible from many orientations (*e.g.,* blemishes, moles, scars). This second strategy, or "transfer by peculiarity," although not as successful as the first one (some faces do not have any localized distinctive marks), would yield performance above chance level. For small rotations, faces with localized distinctive marks could be recognized by using either transfer by configuration or by peculiarity. In this case, subjects may prefer one strategy because it is faster or more accurate or more in agreement with task demands.

The present experiment tests the hypothesis of a dual-strategy by manipulating the presence vs. absence of localized distinctive marks on face images. Performance is measured by a yes/no recognition task. According to the dual-strategy hypothesis, for small rotations (< 30 degrees), subjects base their decision mostly on global configural information; hence, for such rotations, adding a localized distinctive mark should not affect significantly recognition performance. For large rotations (> 30 degrees), the dual-strategy hypothesis states that subjects base their decision on local information. Therefore, adding a distinctive mark should improve significantly the proportion of correct recognitions for large rotations.

The dual-strategy hypothesis does not make explicit predictions about the effect of learned views on recognition performance. However, we can expect that the saliency of a distinctive mark depends on view angles. If this is the case, the usefulness of distinctive marks will depend on the view presented during learning: When the distinctive marks are easy to extract, performance should

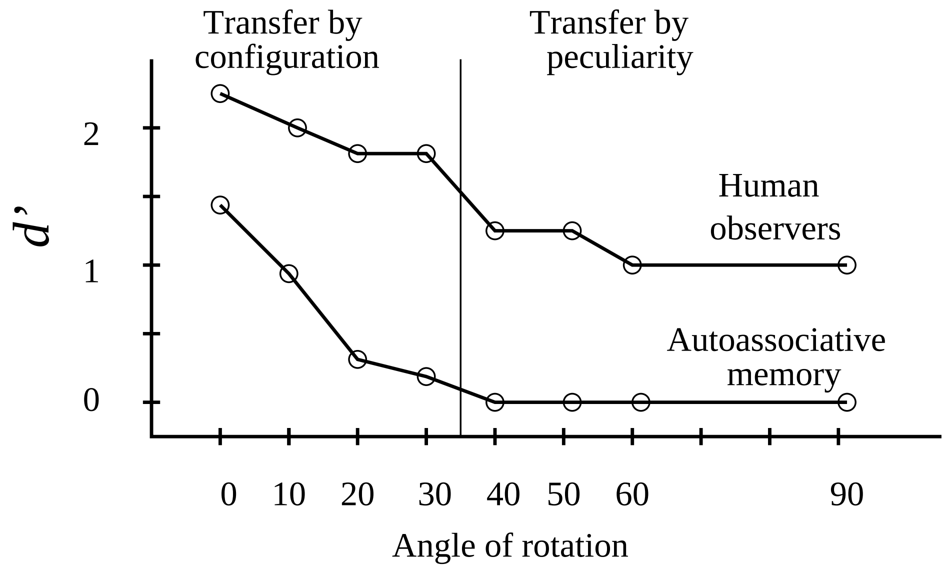

FIGURE 1. Human and neural network performances (as measured by d') on a face recognition task, plotted as a function of the rotation between the view learned and the view tested. According to the dual-strategy model of Valentin *et al.*(in press), the performance for a rotation up to 30 degrees depends upon a "transfer by configuration," whereas the performance for a rotation larger than 30 degrees depends upon a "transfer by peculiarity."

be higher. In addition, some sources report an asymmetry of transfer between frontal, 3/4, and profile views, with transfer from frontal or 3/4 to profile views being better than transfer from profile to frontal or 3/4 views (Hill, Schyns & Akamatsu, 1997; O'Toole, Bülthoff & Walker, 1995; Valentin, 1996). To evaluate a potential relationship between this asymmetry of transfer and the dual-strategy hypothesis we examined if the addition of distinctive marks benefits frontal and profile views differentially.

2. **Experiment**

2.1. *Methods*

Subjects. Twenty-four volunteer students, faculty members, and staff members from the ENSBANA (University of Bourgogne) participated in the experiment. None of the participants was familiar with the faces before the experiment.

Stimuli. Grey-scale pictures of 56 female faces were used as stimuli. Each face was represented by a frontal and a profile view. None had distinctive features such as glasses or jewelry. Half the faces were used as targets and the other half as distractors. Two versions of each view of the target faces were created (*cf.* Figure 2). In the first one, the image was left unchanged, in the second one a distinctive mark was added to the image. Figure 3 displays a sample of the distinctive marks used in the experiment. Note that the distractor faces were left unchanged (*i.e.,* they were never presented with a mark).

Apparatus. The experiment was controlled by a Macintosh Power PC programmed with PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993).

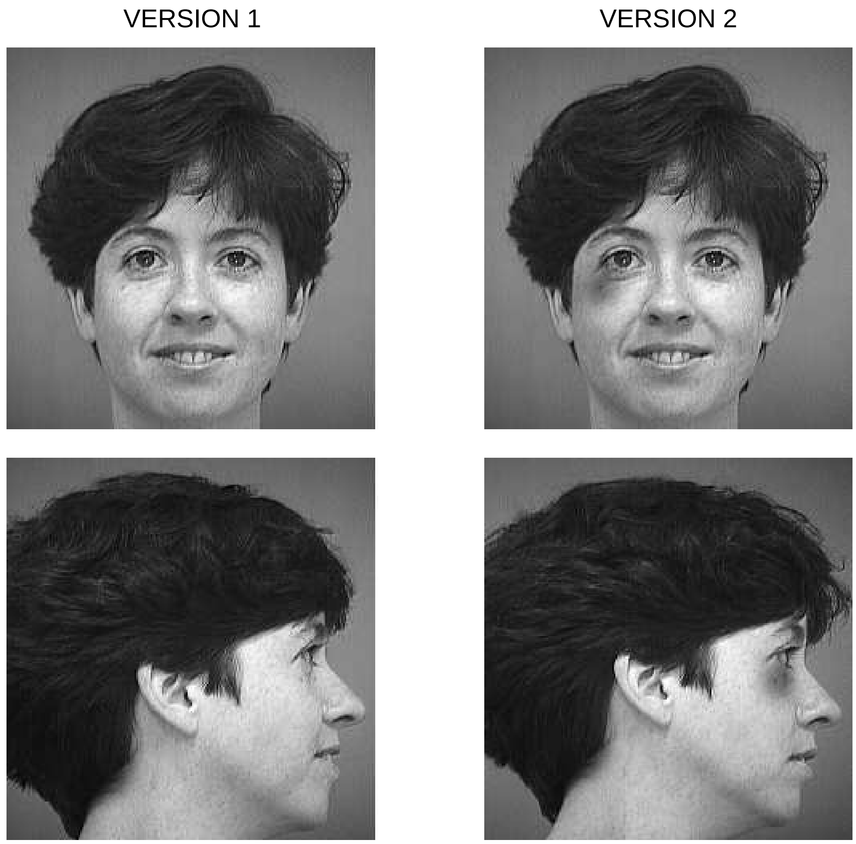

FIGURE 2. The two versions of a target face.

Procedure. The experiment was a standard yes/no recognition task. During the learning phase, subjects were shown 28 faces, each presented on a computer screen for 2 sec, with a 1 sec interstimulus interval. For half the subjects, the faces were presented from a frontal view. For the other half of the subjects, faces were presented from a profile view. In both learning conditions, 14 faces (out of 28) had a distinctive mark added to them. Subjects were asked to memorize the faces. They were informed that a recognition task would follow and that they should be careful to learn the faces but not the pictures. During the testing phase, subjects were shown a second sequence of faces made of the 28 targets and 28 distractors. Half the targets (7 with a distinctive mark, and 7 without a distinctive mark) were presented in the same orientation as during learning and the other half were rotated in depth by 90 degrees. In the same orientation condition, the images were identical to those presented during learning. The subjects were asked to indicate, as quickly as possible, if they recognized the faces. The faces remained on the screen until the subjects indicated their answer. Subjects' response times were also recorded.

A counterbalancing Latin-square procedure was implemented to insure that every target face appeared equally often with and without a distinctive mark as well as in each rotation condition. For both learning and testing, the order of presentation of the faces was randomized and a different order used for each subject.

FIGURE 3. Examples of distinctive marks.

2.2. *Results.* The hit and false alarm rates were calculated for each subject in each condition and combined in a recognition accuracy index ($d' = z_{\text{hit}} - z_{\text{false-alarm}}$) and a decision bias index [$C = -\frac{1}{2}(z_{\text{hit}} + z_{\text{false-alarm}})$]. Hit rates of 100 percent and false-alarm rates of 0 percent were converted respectively to values of $1 - \frac{1}{2N_{\text{Hit}}} = .93$ and $\frac{1}{2N_{\text{FA}}} = .07$, where N_{Hit} (respectively N_{FA}) is the number of faces used to compute the hit rate (respectively false-alarm rate, here $N_{\text{Hit}} = N_{\text{FA}} = 7$; *cf.* Macmillan & Creelman, 1991). This procedure makes the maximum value of d' equal to 2.95.

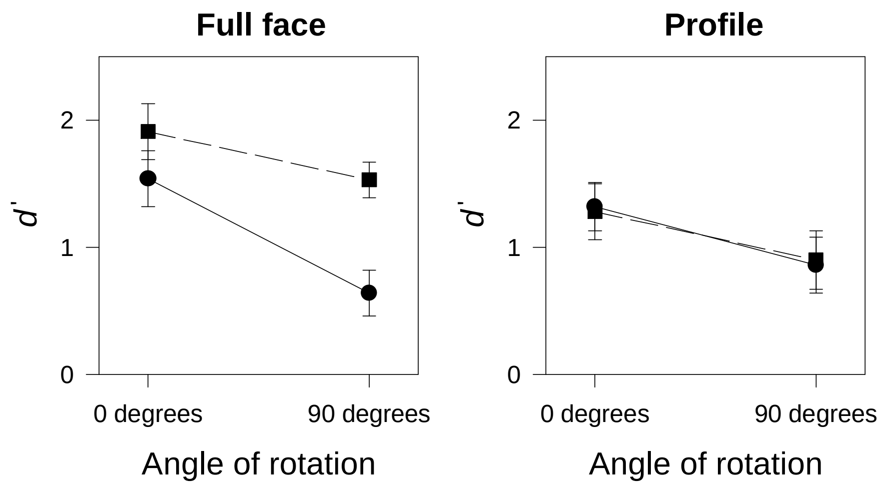

FIGURE 4. Average d' as a function of the rotation angle between learning and test and of the presence vs. absence of distinctive marks when full-faces (left panel) and profiles (right panel) were learned. The squares represent the performance for marked faces and the circles the performance for unmarked faces. Error bars show the standard error of the mean..

Response times, recognition accuracy, and decision bias data were submitted to three separate $2 \times 2 \times 2$ mixed design ANOVAs with *learned view* as a between-subjects factor and *distinctive marks* and *angle of rotation* as within-subjects factors.

Recognition accuracy. The mean d' values are shown in Figure 4. The ANOVA reveals:

- a main effect of the angle of rotation between learning and test, $F(1,22) = 12.55$, $MS_e = .539$, $p < .01$: Recognition performance decreases with the angle of rotation between learning and test (from $d' = 1.51$ to $d' = .98$ on the average).
- a main effect of distinctive marks, $F(1,22) = 9.35$, $MS_e = .256$, $p < .01$: Recognition performance increases when a distinctive mark is present (from $d' = 1.09$ to $d' = 1.41$ on the average).
- an interaction between distinctive marks and type of view presented during learning, $F(1,22) = 9.19$, $MS_e = .256$, $p < .01$. A sub-design analysis conditional on learned view shows a highly significant effect of distinctive marks when a full-face is learned, $F(1,22) = 18.57$, $MS_e = .256$, $p < .001$, but no such effect when a profile is learned. An additional contrast shows that when a full-face is learned, the difference between presence and absence of distinctive marks is significant only in the 90 degrees rotation $F(1,22) = 15.04$, $MS_e = .303$, $p < .001$.

Decision bias. The mean C values are shown in Figure 5. The ANOVA reveals a pattern of results similar to that obtained for d': A main effect of angle of rotation, $F(1,22) = 7.82$, $MS_e = .274$, $p < .05$, a main effect of distinctive marks, $F(1,22) = 9.36$, $MS_e = .064$, $p < .01$, and an interaction between distinctive marks and type of view presented during learning, $F(1,22) = 9.19$, $MS_e = .064$, $p < .01$. On the average, subjects tended to use a stricter criterion when there was a rotation change between learning and test ($C = .30$) than when no change occurred ($C = .01$), and when there was no distinctive marks ($C = .22$) than when a mark was present ($C = .07$). As for d', the effect of distinctive marks is significant only after a 90-degree rotation when full-faces were learned

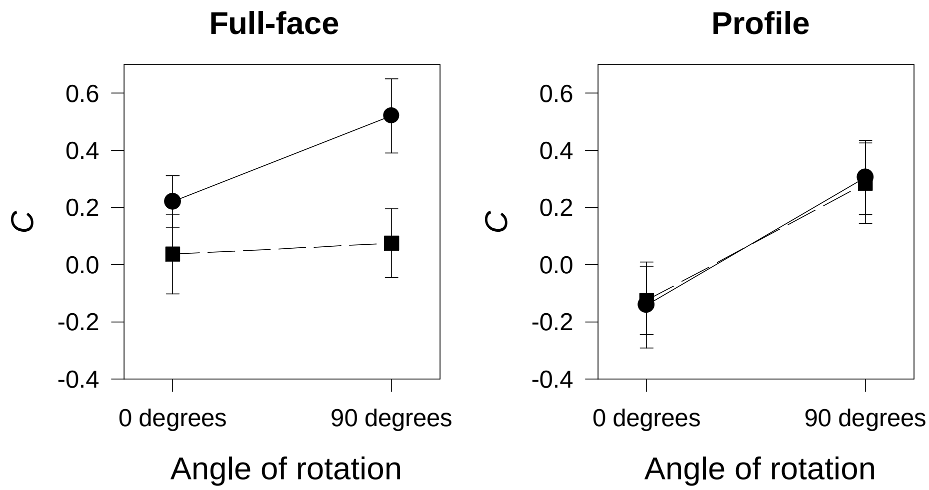

FIGURE 5. Average C as a function of the rotation angle between learning and test and of the presence vs. absence of distinctive marks when full-faces (left panel) and profiles (right panel) were learned. The squares represent the performance for marked faces and the circles the performance for unmarked faces. Error bars show the standard error of the mean.

$F(1,22) = 16.029$, $MS_e = .075$, $p < .001$. No effect of distinctive marks was observed when profiles were learned.

Response time. The median of each subject's target response time distribution was calculated across faces and used as a dependent variable for the following analyses. The average response times (in msec.) are shown in Figure 6 as a function of distinctive marks and angle of rotation. An ANOVA reveals a significant main effect of angle of rotation, $F(1,22) = 9.61$, $MS_e = 791617.3$, $p < .01$ and a close to significant interaction between distinctive marks and type of view presented during learning, $F(1,22) = 3.97$, $MS_e = 590102.10$, $p < .06$. Although this interaction is only marginally significant, it is worth mentioning. In effect, Figure 6 shows that the effect of distinctive marks varies as a function of the view presented during learning. When a full-face was learned, response time decreased in the presence of a distinctive mark. In contrast, when a profile was learned, response time increased in the presence of a distinctive mark.

3. Discussion

The main results of the experiment presented in this paper can be summarized as follows. First, as expected and in agreement with previous work, changing the orientation of faces between learning and test affects subjects' recognition accuracy, response bias and response time. At 90 degrees, recognition accuracy decreases, and both response bias and response time increase. Second, contrary to previous work, we did not find a clear asymmetry of transfer between frontal and profile views on any of the measured dependent variables. The transfer from frontal to profile views is only slightly better than that from profile to frontal views. Third, and most important, we found a strong interaction between distinctive marks and learned views, on all three dependent variables.

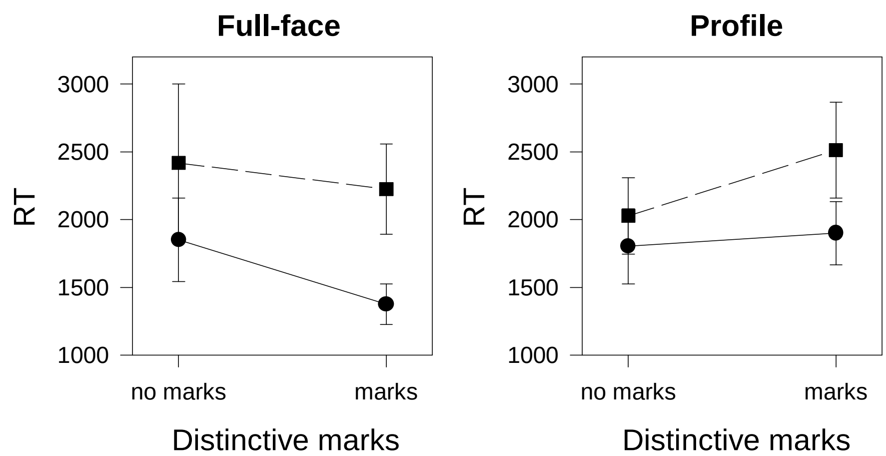

FIGURE 6. Reaction times (in msec.) as a function of the presence vs. absence of distinctive marks (horizontal axis) and of the rotation angle between learning and test when full-faces (left panel) and profiles (right panel) were learned. Note that, here, the squares represent the performance in the 90-degree rotation condition and the circles the performance in the 0-degree rotation condition. Error bars show the standard error of the mean..

When a frontal view was learned, the results agree with the Valentin *et al.*(1998) dual-strategy hypothesis. Recall that, according to this hypothesis, when a small rotation (or no rotation) is applied between learning and test, subjects rely most heavily on general configural information to perform the recognition task. Thus, adding localized distinctive marks to an image should not modify subjects' performance. This is what we observed in the 0-degree condition. On the contrary, when a large rotation is applied, the dual-strategy hypothesis predicts that recognition is based on local information visible from all possible view angles. Thus, adding distinctive marks should improve subjects' performance. Again, this is what we observed in the 90 degrees rotation condition: Subjects were more accurate, faster, and less biased when distinctive marks were added than in the control condition. This pattern of results tends to support the dual-strategy hypothesis. However, nothing in this hypothesis can predict what we observed when a profile view was learned. In effect, even if we consider the fact that profile views have been described in the past as "bad view for transfer" (O'Toole, Bülthoff & Walker, 1995), we would still expect distinctive marks to improve recognition performance.

In our data, distinctive marks do not affect recognition accuracy and error bias when profile views are learned: This is puzzling. A first plausible explanation is, simply, that the distinctive marks are less visible from a profile view than a frontal view. However, an examination of the stimuli shows that even if this is the case for distinctive marks appearing in the central part of the face (*e.g.,* blackened teeth), these marks are few in number. Also some marks (on the side of the face) are, in fact, more visible from a profile than a frontal view. So this explanation alone cannot account for the strong difference in recognition between the learned frontal and profile views.

A second plausible explanation could be that, because they break the symmetry of frontal views, distinctive marks are more salient for these views than for profile views, and thus yield better recognition performance. A direct test of this hypothesis would be to determine if the subjects did actually detect more often the distinctive marks on frontal than on profile views. If this were the case, and if the subjects were consciously aware of the distinctive marks on frontal views, one could argue that the distinctive marks might have been acting as a “label” and as such would have enhanced recognition performance. Although we do not have this direct evidence, an item analysis showed that faces with a symmetry-breaking distinctive mark (*e.g.,* spot on one side of the face) were not better recognized than faces with central or symmetrical distinctive marks (122 hits and 46 misses vs. 114 hits and 54 misses, respectively, $\chi^2 = .914$, $p > .5$). Thus we can think that the difference in disfigurement between profile and frontal views cannot fully explain the observed difference in results between frontal and profile views. Moreover, the fact that the presence of distinctive marks increases the performance only in the 90-degree condition makes somewhat improbable the hypothesis that distinctive marks are playing the role of a label. Indeed, such a hypothesis would predict a performance increase in both the 0 and the 90-degree conditions.

Another, and more probable, explanation would be that subjects use a different type of information, or process, depending on the type of view they are asked to memorize. For a frontal view, they would pay more attention to the texture of the face, and thus would detect the distinctive marks. For a profile view, they would pay more attention to the general shape of the face (hair and silhouette), and thus would not detect textured distinctive marks. In agreement with this hypothesis, in the 90-degree condition, the response time decreases with the presence of distinctive marks when a frontal view was learned, but it increases when a profile was learned. If distinctive marks are not encoded when a profile view is learned, their presence at test when a frontal view is presented may not help, and may even “distract” the subjects. Moreover, a shift of attention from texture to shape could also explain the asymmetry of transfer generally observed between frontal and profile views. Texture information might, indeed, be easier to transfer, or be less variable from one view to another, than shape information.

In conclusion, the answer to the question “how do we transfer information from a single view of a face to new views of this face?” is: It depends on the view and the amount of transfer. Our results suggest that, when asked to memorize a frontal view of a face, we focus our attention mostly on textural information. This type of information will allow us to use a global matching strategy, or “transfer by configuration,” when the angle of rotation between the to-be-recognized face and the learned face is not too large and to rely on localized information, or distinctive marks (“transfer by peculiarity”) otherwise. On the other hand, it seems that when asked to memorize a profile view of a face we focus our attention, mostly, on shape information. This type of information, also slightly less efficient than the texture information, allows for a “transfer by configuration” but not for a “transfer by peculiarity.” Indeed, a peculiarity corresponds, in this context, to a localized area of abnormal texture. Further work, however, is needed to look more closely at the transfer of information from a profile view of a face. Do subjects try to predict the shape from another view angle using some transformation rules? Or do they simply try to match the two views? It would also be of interest to investigate the effect of distinctive marks for intermediate views between frontal and profile views. Such an investigation could serve to determine if the effect of localized marks declines gradually or abruptly across learning views.

Acknowledgements. Thanks are due to two anonymous reviewers for helpful comments on an earlier draft of this paper. Correspondence should be sent to: Dominique Valentin, ENSBANA, 1, Esplanade Erasme, Campus Universitaire, 21000 Dijon, France, E-mail: `valentin@u-bourgogne.fr`, tel: (33) 3 80 39 66 48 or Hervé Abdi, School of Human Development, The University of Texas at Dallas, MS: Gr4.1, Richardson, TX 75083-0688, E-mail: `herve@utdallas.edu`, tel: (1) 972 883 20 65.

References

[1] Abdi, H. (1987). A generalized approach for connectionist auto-associative memories: Interpretation, implication, and illustration for face processing. In J. Demongeot (Ed.) *Artificial intelligence and cognitive science.* Manchester: Manchester University Press.

[2] Abdi, H., Valentin, V., Edelman B., (1999). *Neural Networks.* Thousand Oaks (CA): Sage.

[3] Baddeley, A. and Woodhead, M. (1981). Techniques for improving eyewitness identification skills. Paper presented at the SSRC law and psychology conference Trinity college Oxford.

[4] Bruce, V. (1982). Changing faces: Visual and non-visual coding process in face recognition. *British Journal of Psychology*, 73:105–116.

[5] Cohen, J.D., MacWhinney, B., Flatt, M., and Provost, J. (1993) PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods Instruments & Computers,* 25:257–271.

[6] Davies, G., Ellis, H., and Shepherd, J. (1978). Face recognition accuracy as a function of mode of presentation. *Journal of Applied Psychology*, 63:180–187.

[7] Hill, H., Schyns, P. and Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, 62:201–222.

[8] Krouse, F. (1981). Effects of pose, pose change, and delay on face recognition performance. *Journal of Applied Psychology*, 66:651–654.

[9] Logie, R., Baddeley, A., and Woodhead, M. (1987). Face recognition, pose and ecological validity. *Applied Cognitive Psychology*, 1:53–69.

[10] Macmillan, N. and Creelman, C. (1991). *Detection theory: A user's guide.* Cambridge University Press, Cambridge.

[11] O'Toole, A., Bülthoff, H., Troje, N., and Vetter, T. (1995). Face recognition across large viewpoint changes. In Bichsel, M. *Proceedings of the International Workshop on Automatic Face and Gesture Recognition*, University of Zurich. MultiMedia Laboratory, Department of Computer Sciences.

[12] O'Toole, A., Bülthoff, H., and Walker, C. (1995). Face recognition across viewpoint. TR 21, Max-Plant-Institut für biologishe Kybernetik.

[13] Patterson, K. and Baddeley, A. (1977). When face recognition fails. *Journal of Experimental Psychology: Human Learning and Memory*, 3:406–417.

[14] Valentin, D. (1996). *How come when you turn your head I still know who you are: Evidence from computational simulations and human behavior.* Unpublished doctoral dissertation, University of Texas, Dallas.

[15] Valentin, D., Abdi, H., and Edelman, B. (1997). What represents a face: a computational approach for the integration of physiological and psychological data. *Perception*, 26:1271–1288.

[16] Valentin, D., Abdi, H., Edelman, B. and Posamentier, M. (in press). 2D or not 2D? That is the question: What can we learn from computational models operating on 2D representation of faces. In T. Wenger and J. Townsend (Eds.) *Computational, geometric, and process perspectives on facial cognition: context and challenges.* Hillsdale, MA: Erlbaum.