THE ROYAL SOCIETY

doi 10.1098/rspb.2000.1274

The Noh mask effect: vertical viewpoint dependence of facial expression perception

Michael J. Lyons[1*], Ruth Campbell[2], Andre Plante[1], Mike Coleman[2], Miyuki Kamachi[3] and Shigeru Akamatsu[3]

[1]*Advanced Telecommunications Research International Media Integration and Communications Research Laboratories and*
[3]*Advanced Telecommunications Research International Human Information Processing Research Laboratories, 2-2-2 Hikari-dai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan*
[2]*Department of Human Communication Science, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PG, UK*

Full-face masks, worn by skilled actors in the Noh tradition, can induce a variety of perceived expressions with changes in head orientation. Out-of-plane rotation of the head changes the two-dimensional image characteristics of the face which viewers may misinterpret as non-rigid changes due to muscle action. Three experiments with Japanese and British viewers explored this effect. Experiment 1 confirmed a systematic relationship between vertical angle of view of a Noh mask and judged affect. A forward tilted mask was more often judged happy, and one backward tilted more often judged sad. This effect was moderated by culture. Japanese viewers ascribed happiness to the mask at greater degrees of backward tilt with a reversal towards sadness at extreme forward angles. Cropping the facial image of chin and upper head contour reduced the forward-tilt reversal. Finally, the relationship between head tilt and affect was replicated with a laser-scanned human face image, but with no cultural effect. Vertical orientation of the head changes the apparent disposition of facial features and viewers respond systematically to these changes. Culture moderates this effect, and we discuss how perceptual strategies for ascribing expression to familiar and unfamiliar images may account for the differences.

Keywords: face perception; facial expression perception; Noh masks

1. INTRODUCTION

The perception of affect is delivered by a variety of cues. Body posture is one: positive affect is accompanied by an upright posture with the head held high, or even thrown back. Some states of negative affect (sadness, fear) are signalled by a bowed head and a more crouched posture (Darwin 1872). The face, however, offers the most salient perceptual cues to emotional state. The actions of the facial muscles generate characteristic changes in shape and position of the facial features, especially the brows and mouth (Ekman & Friesen 1978). Detection of such changes by the perceiver's visual system requires sensitivity to the fine metric properties of the facial surface. Rigid-body transformation of the head, specifically rotation out of plane, as with vertical movements of the head, distorts configural relationships on the face as they appear in the two-dimensional (2D) projection of the visual world on the retina. The consequences of this are paradoxical. For body posture signals, positive affect judgement should vary positively with degree of upward head tilt. But, under full frontal viewing, upward head tilt projected into two dimensions can generate image transformations which reduce upward curvature of the mouth, and change the angle and disposition of other facial features including the brows. This may give the impression of a more negative expression. A similar paradox occurs for a forward tilted head. To what extent are viewers sensitive to these conflicting influences, and how may they resolve—or fail to resolve—them?

Our interest in this question was stimulated when we learned of an illusion of facial expression perception involving masks used in Japanese traditional Noh drama (Komparu 1983). Certain Noh masks, particularly those used to portray young female roles (including the masks known as Ko-omote, Wakaonna, and Magojiro), appear to change expression as the vertical inclination of the mask changes (see figure 1). Tilt the mask forward and it appears to smile; tilt it backwards and it appears sad.

The lifelike changeability of the masks has been known in the Noh theatre for centuries (the earliest such masks were carved in the Kamakura period (1192–1333)) and is considered an important ingredient in the mysterious atmosphere of Noh drama.

We thought it possible that one aspect of the actor's skill, and that of the mask maker, could be to enhance the perceptual effects outlined above. These suggest that small changes in vertical orientation of the mask could, under appropriate viewing conditions, give rise to the illusion that the face has taken different expressions.

Three experiments were conducted to test the perception of the mask systematically. The following questions were posed: (i) Does the image of a Noh mask generate different perceptions of affect as a function of its vertical orientation? (ii) To what extent does culture moderate any effects? In particular, does familiarity with the Noh mask image (Japanese) generate different functions than those found for cultures with no such familiarity (British). (iii) Do effects of inclination generalize to other faces?

*Author for correspondence (mylons@mic.atr.co.jp, http://www.mic.atr.co.jp/~mlyons).

Received 13 June 2000 *Accepted* 21 August 2000

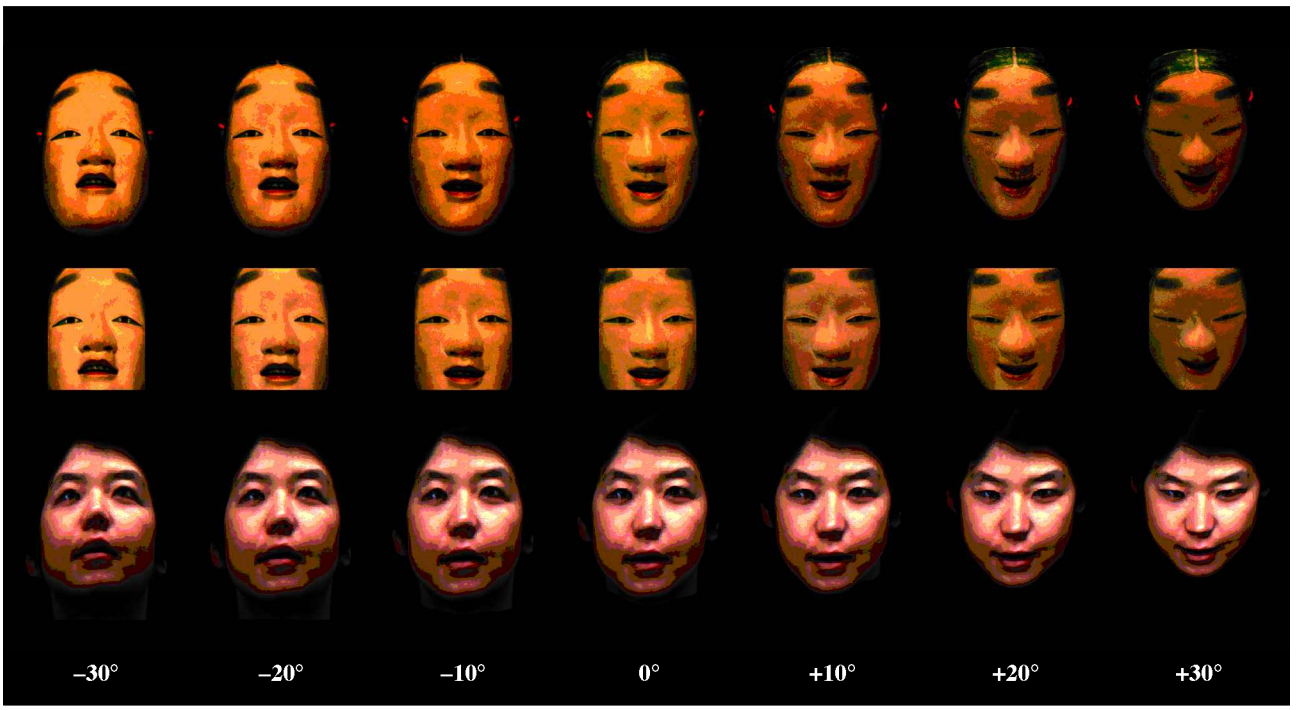

Figure 1. Top row, examples of the stimuli used for experiment 1; full Magojiro Noh mask at different angles of inclination. Middle row, examples of the stimuli used for experiment 2; same as the above, but the sides, top and chin areas have been cropped. Bottom row, examples of the stimuli used for experiment 3; images rendered from a 3D laser scan of a Japanese female head.

2. MATERIAL AND METHODS

(a) *Stimuli*

(i) *Experiment 1*

Stimuli for experiment 1 consisted of photographs of a Noh mask (figure 1) at 13 inclinations, from −30° to +30° in equal 5° increments. An antique Magojiro mask, used for young female roles, dating to the Edo period (1600–1868) was photographed on a Noh stage under lighting conditions similar to those used during a performance (tungsten illumination over the stage with some ambient illumination from the room). The tripod-mounted mask was photographed from a frontal viewpoint using a tripod-mounted digital camera (Kodak Professional DCS 460; Eastman Kodak Co., Rochester, NY, USA) from a distance of 7.7 m with a 200 mm lens. The 3060×2036 pixel 24-bit colour images were cropped and re-sampled to 300×400 pixel TIFF images.

(ii) *Experiment 2*

The stimuli for experiment 2 comprised the same images, but were cropped to reveal only the internal features of the face.

(iii) *Experiment 3*

Stimuli for experiment 3 were derived from the head and face of a 30-year-old Japanese female model posing a neutral expression somewhat similar to that of the Noh mask. A Cyberware 3030 RGB/PS colour three-dimensional (3D) scanner was used to acquire shape and colour information of the model's head. The 24-bit RGB colour map was acquired under room light from an overhead fluorescent lamp. Spatial resolution of the scan, in cylindrical coordinates, is 0.7 mm in the vertical axis and $2\pi/512$ radians for the longitudinal angle. The resolution of the radius depends on local surface reflectance properties and is estimated at *ca.* 1 mm for skin areas. A modified version of the Cyberware Echo software was used to reconstruct a 3D model of the head (Cyberware, Inc., Monterey, CA, USA). Screen captures were taken at 13 (virtual) head inclinations, the face orientated frontally and saved as 24-bit 300×400 pixel TIFF images. The vertical viewing angles varied from −30° to +30° in equal 5° increments. Interocular distance and eye position were normalized for each stimulus set and matched across sets. Shape-only models of the mask (figure 2*a*,*b*) and the human face (figure 2*b*) were rendered from 3D laser scans using a Gouraud shading model with Cyberware Echo software. The mask used in figure 2 is a Ko-omote mask (similar to Magojiro) belonging to one of the authors.

(b) *Procedure*

Experiments were run in separate laboratories in London and Kyoto. In each case the stimuli were displayed on a 17 in 24-bit colour computer monitor in a slightly darkened room. Viewing distance was *ca.* 60 cm. Following a practice trial, four epochs of all 13 stimuli were presented in succession, with presentation order randomized within each epoch. Presentation order was as follows: fixation point (500 ms)—blank (400 ms)—stimulus (300 ms). Subjects were instructed to respond whether the stimulus face appeared happy or sad by pressing the left or right shift key. Japanese subjects were instructed in Japanese using the terms '*yorokobi*' and '*kanashimi*'. Left or right assignment of response keys was counterbalanced across subjects. The words 'happy' and 'sad' (in English for both subject groups) appeared on the response-appropriate side of the screen for each subject to maintain correct response orientation. All Japanese

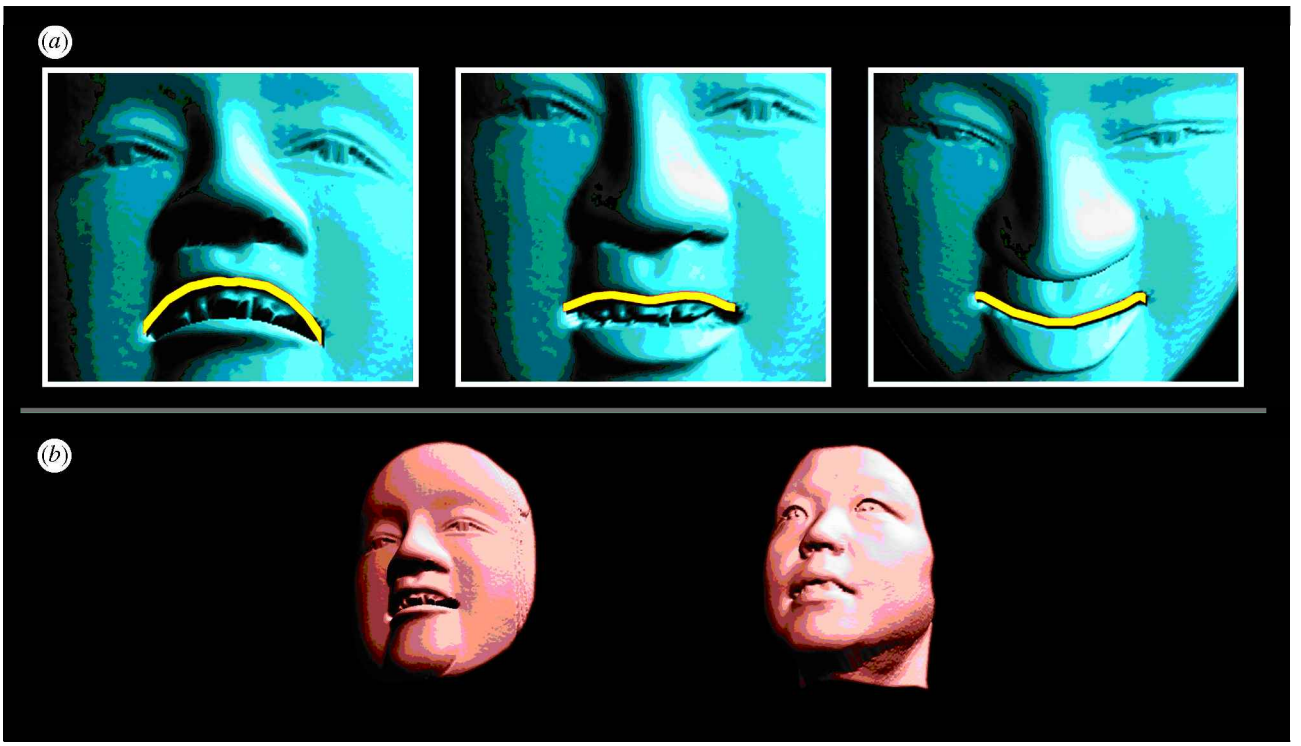

Figure 2. (a) Images rendered from a 3D laser scan of a Noh mask at three inclinations. The contour of the top lip has been highlighted, drawing attention to the variation of the 2D projection of this feature with inclination. The projection of other mask features is also affected. (b) Rendered using shape-only information from 3D laser scans of a Noh mask (left) and Japanese female face (right).

subjects were familiar with the English terms. Reaction times and decision type were recorded automatically for each subject for each trial. The brief exposure time and central fixation cue were used to control the visual information received by each subject. With free viewing conditions, subjects would spend variable amounts of time looking at different areas of the face. With the paradigm used here, we measure the impression received after a glance at the stimulus.

(c) *Subjects*

Different subjects were run for each of the three experiments. There were ten females and ten males from each cultural group for each experiment, making 120 subjects in total. Subjects were undergraduates, graduates and staff from Doshisha University, Kyoto, and University College London. Ages ranged from 18 to 50 years. All had normal or corrected-to-normal vision. All were either native to the country of testing or had first-school-education in that country. The Japanese subjects were familiar with Noh masks as images or, occasionally, objects. None of the UK subjects had familiarity with Noh or had visited Japan.

3. RESULTS

All three experiments had the same mixed, three-factor repeated-measures design. There were 13 levels of the first factor (inclination), which varied within subjects, and one level of the first between-subjects factor (culture) and the second between-subjects factor (gender).

The binomial response (happy or sad) is at the nominal level of measurement. The proportion of 'happy' responses calculated over the four trials, r, was transformed using the equation $r' = \arcsin(\sqrt{r})$. This procedure allows treatment of the response proportions at the interval level (Johnson *et al.* 1992). Except where described otherwise (see § 3(a,b)), histograms of the response proportions showed unimodal distributions. Reaction times (medians) were also examined but did not show any systematic relationship to the other variables and are not reported here. Gender had no effect on any of the analyses and effects of gender were not considered further. Table 1 outlines the significant finding for each experiment. The relevant graphs are shown in figure 3. These analyses show a significant relationship between angle of inclination and proportion of 'happy' responses in all three experiments. The Noh mask, but not the scanned face, is classified differently by Japanese and British viewers. Further analyses and their justification in terms of individual experimental hypotheses are reported in § 3(a–c).

(a) *Experiment 1*

The experimental hypothesis was that Noh mask images would generate systematic changes in perceived affect as a function of angle of inclination. The perception of facial expressions is thought to be similar for our two cultural groups, though lower recognition rates for negative emotions have been reported for Japanese viewers (Matsumoto 1992). Moreover, the expression perceived in cultural artifacts such as Noh masks may require some familiarity with the mask or with culturally stylized expressions.

The results confirmed the predictions in general terms, but with some important deviations (see figure 3*a*). One is that the groups differed in the inclination at which the proportion of 'happy' responses reaches 50%, Japanese

Table 1 *Summary of F-values, separate ANOVAs for each experiment.*

(Calculated using the SPSS GLM procedure.)

	$F_{12,432}$ main effect of inclination	$F_{1,36}$ main effect of group	$F_{12,432}$ group× inclination interaction
experiment 1 full mask	20.6 $p < 0.001$	4.19 $p < 0.05$	6.55 $p < 0.001$
experiment 2 cropped mask	39.7 $p < 0.001$	n.s.	3.21 $p < 0.05$
experiment 3 scanned face	19.6 $p < 0.001$	n.s.	n.s.

viewers reaching the crossing point at more negative angles than British viewers. This may reflect different boundaries in terms of perceived facial expression on the categories of 'happy' and 'sad' and their cognates in Japanese, or lower rates of sadness recognition for Japanese viewers. This issue is addressed further in experiment 3.

A second unexpected finding is that the groups differed in the pattern of the relationship with inclination angle (interaction of group and inclination was highly significant). While the relationship was monotonic over this range for the British subjects, for the Japanese subjects, the proportion of 'happy' responses peaked at 5° and then dipped. At 30°, the proportion of 'happy' responses was less than 50%. Examination of the histogram of response proportions (figure 4) showed significant bimodality in the Japanese viewers responses beyond 15°.

Why should this change in perceived expression occur? One possibility is suggested by the postural cues in the images of the head. A head bowed forwards suggests 'sadness' and the two groups may weight the posture and internal features cues differently. The conflict between the posture and internal feature cues may explain the response variability in the Japanese group. For the Noh mask images used, pose cues are most visible in the disposition of the top of the head and the chin with change in inclination of the head.

(b) *Experiment 2*

In this experiment, the face images were cropped to diminish cues to head pose and emphasize internal features of the face. The experimental prediction was that this may eliminate the dip in the percentage of 'happy' responses for Japanese viewers as the mask is tilted forward. The results supported this (figure 3*b*). In this study, the 'dip' at greater positive inclinations was reduced, though not completely eliminated. Thus, it appears that Japanese viewers take account of cues to head pose in ascribing expression to the image of the vertically inclined mask. Response histograms of the Japanese viewers exhibited some bimodality for these angles (figure 4), suggesting the conflict between featural and residual pose cues is not completely resolved.

Otherwise, this study replicates the main findings of experiment 1. Once more, a group difference emerged between Japanese and British viewers, in the same direction as that of experiment 1.

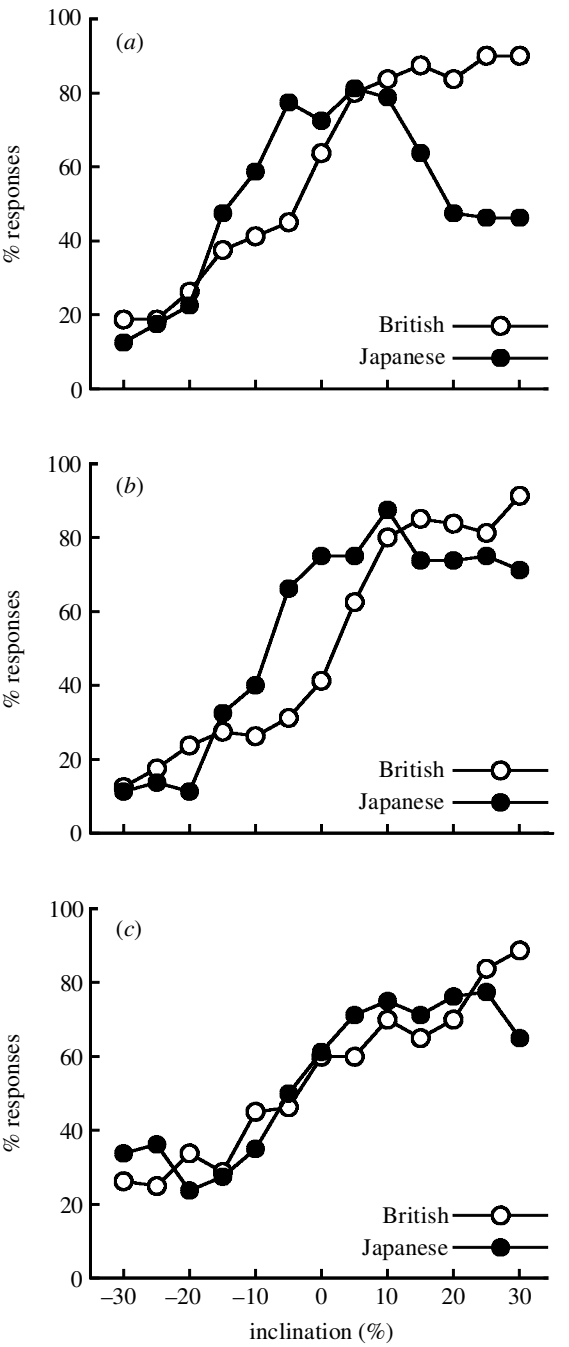

Figure 3. Proportion of 'happy' responses, coded as percentage of total responses, versus inclination angle for (*a*) experiment 1, full mask; (*b*) experiment 2, cut mask; and (*c*) experiment 3, scanned face.

One reason for this may be that the terms 'happy' and 'sad' in English and Japanese may not share similar extensions. This would suggest that Japanese may be more willing than British viewers to ascribe 'happy' (a socially acceptable facial signal) to a relatively 'unhappy' face. If this were the case, we would expect a similar disparity between groups to emerge when images of natural faces are perceived. Experiment 3 explores this possibility.

(c) *Experiment 3*

This experiment used stimuli generated from a 3D laser scan of a human face to explore the question: Do group

Table 2. *Approximate facial action unit coding of Noh mask stimuli*

(Codes for facial actions are AU1, inner brow raise; AU2, outer brow raise; AU6, cheek raise; AU10, upper lip raise; AU12, lip corner pull; AU15, lip corner depress; AU26, jaw drop; AU41, lids droop (Ekman & Friesen 1978). (+), possible presence of action; +, definite presence of action pattern. Letters indicate intensity: a, present to some extent; d, most marked.)

mask inclination	AU1	AU2	AU6	AU10	AU12	AU15	AU26	AU41
−30°	c	—	—	a	—	a	a	b
30°	—	c	+	(+)	—	d	—	a

differences in ascribing expression to a cultural artifact, the Noh mask, extend to natural face images? If they do, we may infer that cultural and linguistic interpretations of facial expression may differ between these groups. If they do not, then the Noh mask may have special perceptual status for Japanese viewers. The findings support the latter conclusion. The relationship between inclination angle and happy–sad judgements was nearly identical in both groups (figure 3*c*). *t*-tests exploring group differences at each orientation point reached significance only at 30° ($p < 0.05$, d.f. = 38), where a slight dip is observed for the Japanese subjects.

We can conclude that both Noh mask effects, the dip in the function at high angles of forward tilt in experiment 1 and the 'happier' classification at most other angles in experiments 1 and 2, reflected a cultural phenomenon—but one related to *perceptual* processing differences between the groups.

Though the laser-scanned face did not replicate the lighting conditions of the naturally photographed images used in experiments 1 and 2, the relationship between inclination and judged expression still held, suggesting that the difference in lighting differences did not contribute notably to the illusion for this set of conditions.

4. DISCUSSION

Three experiments confirmed that the angle of vertical inclination of a face profoundly influences a simple expression discrimination task: faces tilted down have a happier cast than those tilted back. This may be understood in terms of the projection of the 3D facial surface onto the image (see figure 2*a*). The basic phenomenon is related to the widely known trick in which the expression of a face on a folded banknote may be varied by tilting the plane of the note. Cavanagh *et al.* (1988) noted the effect as an example of the failure of shape constancy under rotation in depth. We know of only one previous experimental study of the effect of vertical viewpoint on facial expression perception (Kappas *et al.* 1994). Kappas' experiments used video clips of posed dynamic expressions and a schematic wire-frame model of the face, both quite different from the Noh mask stimuli and scanned face used here. In addition, that work did not involve pose variations as in our experiments 1 and 2, or study different cultural groups.

The facial action coding system (FACS), developed by Ekman & Friesen (1978), is a general method for describing visible changes to the face due to muscular action. FACS codings of the two images of the Noh mask, at extreme forward and backward tilt, are reported in table 2. The FACS method relies partially on the observation of creasing of the skin that results from facial actions; the Noh mask is smooth, hard, and does not wrinkle. Moreover, the mask depicts shaved eyebrows and heavy brow marks painted on the upper forehead, a cosmetic fashion of the Heian period. Therefore, the interpretation of the apparent changes in feature shape as action units should be considered as an approximate application of FACS. These caveats aside, comparison of table 2 with FACS codes for facial expression prototypes as reported in Ekman & Friesen (1978) and Benson *et al.* (1999) is instructive in understanding the Noh mask effect.

The forward tilted mask suggests the action of AU6 (cheek raise) and AU12 (lip corner pull), both key elements of the prototypical happy expression. AU26 (jaw drop) is also present, but does not interfere with the impression given by the other units. The action of AU2 (outer brow raise) might be expected to make the expression seem less happy; however, the impact of brow action is reduced by the unusual position of the brow marks. The sad cast of the backward tilted mask seems to be conveyed mainly by the presence of AU1 (inner brow raise) and AU15 (lip corner depress) and perhaps AU41 (lids droop). AU1 usually occurs with AU4 in the prototypical sad expression, but the unusual position of the brows may affect the viewer's interpretation of their configuration.

A surprising but consistent finding was that the Noh mask elicited different responses in the two cultural groups. Most strikingly, in experiment 1, the Japanese responses peaked early then dipped strongly with increasing angle, and bimodal response distributions were observed beyond tilt angles of 15°. Comparison of experiments 1 and 2 (which emphasized the internal mask features and reduced cues to mask pose) suggests a possible explanation for the latter cultural difference: the Japanese viewers may place more importance on postural cues in the foward tilted mask than the British viewers, and these conflict with the changes visible in the internal features. British viewers may have based their decisions solely on curvature of the lips, an internal feature which reliably signals 'happy–sad' (Calder *et al.* 2000) and which varies monotonically with head inclination.

A second cultural effect was present in both experiments 1 and 2: Japanese viewers ascribed happiness to the mask at greater negative angles of tilt. Absence of the effect for the scanned face (experiment 3) implies that it was not caused by differences in linguistic boundaries of happiness and sadness, but may have resulted from different perceptual strategies for processing the internal features of the mask. A salient cultural difference between the two groups is that the Noh mask was unfamiliar to the British group, but familiar to the Japanese. The Noh mask occasionally appears in the Japanese media, though

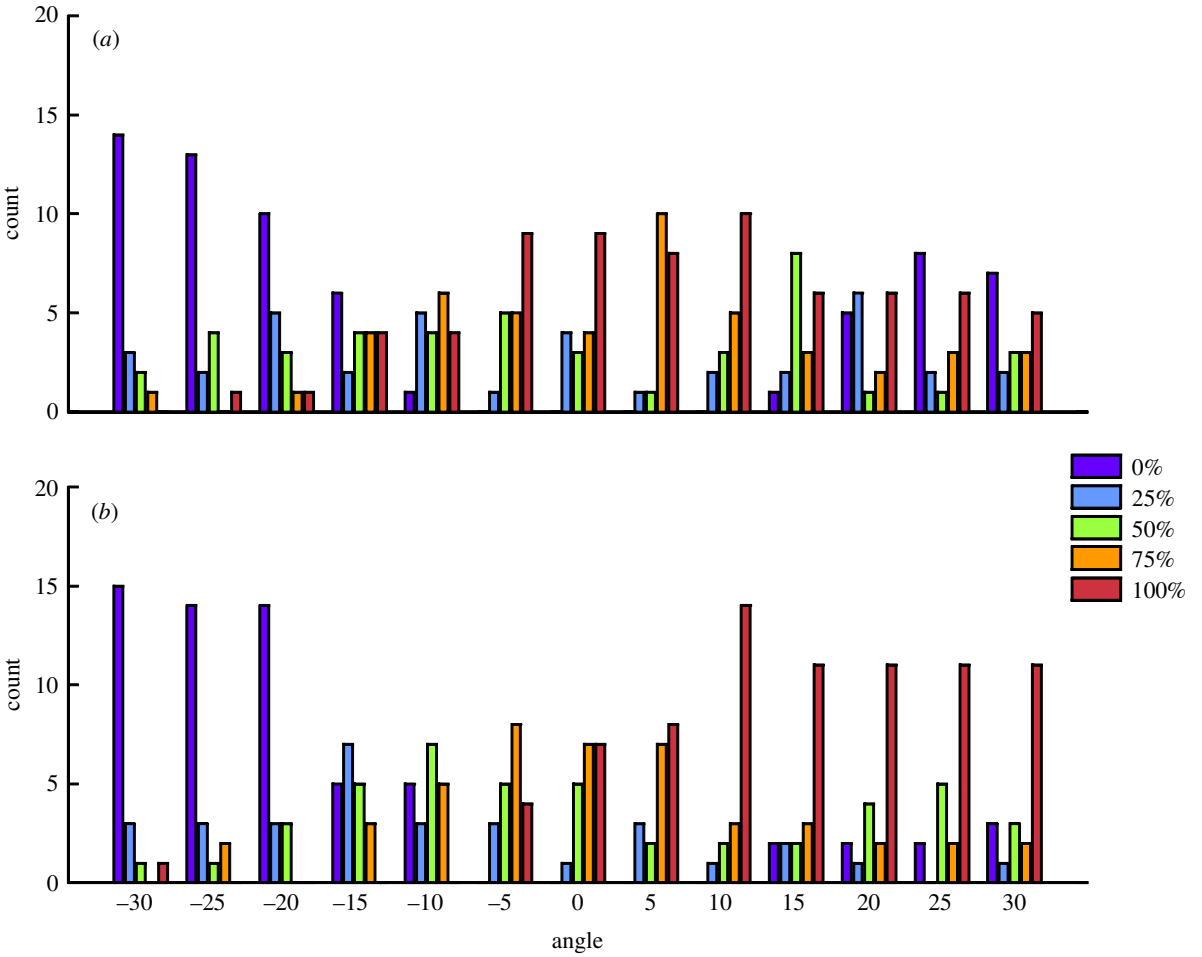

Figure 4. Histograms of proportion of 'happy' responses versus inclination angle for Japanese subjects: (*a*) experiment 1, full mask; and (*b*) experiment 2, cut mask.

an understanding of, or interest in, Noh as a tradition is no longer widespread in the general population. The skilled processing of faces has typically been described as configurational (Diamond & Carey 1986; Young *et al.* 1987). That is, skilled viewers take account of the various face features and their disposition in coming to a unified account of the identity or reading of the face. Their reading of the face cannot be predicted on the basis of local featural detail. One possibility is that familiarity may have delivered a greater degree of configural processing for the mask in Japanese than in British viewers. Only further experiments will indicate what facial aspects are used by the different cultural groups in making their judgements.

At the outset of these studies, we speculated that the 3D structure of the Noh mask and the disposition of the painted features may be intentionally designed to elicit changes of perceived expression with small changes in pose. Examination of the 3D structure of the mask showed, for example, that the depth of the mouth region is exaggerated relative to the human face. Our psychological studies confirmed that small changes in pose of the mask lead to significant changes in perceived affect. A forward tilted mask appeared relatively happy and one tilted backwards, relatively sad. Paradoxically, however, in the stylized use of mask pose in Noh drama, the convention is the opposite to our findings. In one gesture known as *terasu* (shining), signifying a happy state, the mask is turned upwards. In another known as *kumorasu* (clouding), signifying a sad state, the mask is turned downwards (Komparu 1983).

In this connection, it is notable that Zeami (1363–1443), the most influential early Noh dramatist, ranked *yugen*, or subtle profundity, as the highest aesthetic principle of Noh (Zeami 1968). In the framework of the Noh world, a joyful pose tempered with a slightly sad mouth may be appreciated as more beautiful than a direct expression of joy. Likewise, sadness or pain masked with a smiling mouth suggests more emotional complexity than a display of pure sadness.

A further interpretation is possible, not necessarily in conflict with the above. The psychometric curves (figure 3) show that small changes in inclination angle significantly affect perceived facial expression. Minor movements of the actor's head may trick viewers into thinking that the internal features of the mask are moving non-rigidly as if it were an animated living face. One of the authors (M.J.L) has observed this effect while watching a Noh play. In a related perceptual effect, a rigid 3D stick man figure rocked longitudinally back and forth can appear to walk with non-rigid limb movement in 2D projection (Sinha & Poggio 1996). Another perceptual trick, the hollow mask illusion (Gregory 1973; Hill & Bruce 1996) demonstrates that the visual system has *a*

priori knowledge of the depth profile of faces. It is interesting to speculate that rigidity failure in face perception may be enhanced by increasing the curvature of certain features, such as the mouth, beyond what the human visual system expects. This may underlie the exaggeration of the mouth region in the design of the Noh mask. Full investigation of this hypothesis will require experiments using dynamic stimuli.

After this study was completed, we discovered a recently published study (Minoshita *et al.* 1999) on the effect of vertical inclination on facial expression perception in Noh masks. The findings of that study confirm our conclusion that the apparent affect of the mask changes significantly with viewing angle. The experimental design of that study, however, used strongly asymmetrical lighting on the mask, a more extreme range of vertical angles, free viewing conditions, and a very large number of emotional descriptors. It is therefore difficult to compare the results in any detail.

We are grateful to Hotaka Komparu for allowing us to photograph the Magojiro mask and for a very helpful discussion about Noh drama; the management at Nara Shin Kokai Do for access to the Noh stage; the Yokoya laboratory at the Nara Institute of Science and Technology for use of the Cyberware scanner; Akiko Tohma and Katsunori Isono for their generous help; Larry Maloney, Hiroshi Yamada, and Jiro Gyoba for useful discussions; Patrick Cavanagh for sharing his unpublished notes; and Vicki Bruce for encouraging us to pursue this study.

REFERENCES

Benson, P. J., Cambell, R., Harris, T., Frank, M. G. & Tovée, M. J. 1999 Enhancing images of facial expressions. *Percept. Psychophys.* **61**, 259–274.

Calder, A. J., Young, A. W., Keane, J. & Dean, M. 2000 Configural information in facial perception. *J. Exp. Psychol. Hum. Percept. Perf.* **26**, 527–551.

Cavanagh, P., Peters, S. & Von Grnau, M. 1988 Rigidity failure and its effect on the Queen. *Perception* **17** (Suppl.), 27A.

Darwin, C. 1872 *The expression of the emotions in man and animals*. London: John Murray.

Diamond, R. & Carey, S. 1986 Why faces are and are not special: an effect of expertise. *J. Exp. Psychol.* **115**, 107–117.

Ekman, P. & Friesen, W. V. 1978 *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Gregory, R. L. 1973 The confounded eye. In *Illusion in nature and art* (ed. R. L. Gregory & E. H. Gombrich), pp. 49–96. London: Duckworth.

Hill, H. & Bruce, V. 1996 A comparison between the hollow face and hollow potato illusions. *Perception* **23**, 1335–1337.

Johnson, N. L., Kotz, S. & Kemp, A. W. 1992 *Univariate discrete distributions*, 2nd edn. New York: Wiley.

Kappas, A., Hess, U., Barr, C. & Kleck, R. 1994 Angle of regard: the effect of vertical viewing angle on the perception of facial expressions. *J. Nonverbal Behav.* **18**, 263–280.

Komparu, K. 1983 *The Noh theatre: principles and perspectives*. New York and Tokyo: Weatherhill/Tankosha.

Matsumoto, D. 1992 American–Japanese cultural differences in the recognition of universal facial expressions. *J. Cross-Cult. Psychol.* **23**, 72–84.

Minoshita, S., Satoh, S., Morita, N., Tagawa, A. & Kikuchi, T. 1999 The Noh mask test for analysis of recognition of facial expression. *Psychol. Clin. Neurosci.* **53**, 83–89.

Sinha, P. & Poggio, T. 1996 Role of learning in three-dimensional form perception. *Nature* **384**, 460–463.

Young, A. W., Hellawell, D. & Hay, D. C. 1987 Configural information in face perception. *Perception* **16**, 747–759.

Zeami, M. 1968 *Kadensho*. Kyoto, Japan: Doshisha University Sumiya-Shinobe Publishing Institute.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.