

DEVELOPMENTAL CHARACTERISTICS IN JUDGMENTS OF EMOTION FROM NONVERBAL VOCAL CUES

DAVID MATSUMOTO

University of California, Berkeley

and

HIROMU KISHIMOTO

Meiji University, Tokyo, Japan

ABSTRACT. In an exploratory study on the developmental trends in judgments of emotion from nonverbal vocal cues, content free speech tapes designed to express one of the four emotions of happiness, surprise, sadness, and anger were presented to 47 American and 56 Japanese children, ages 4 through 9. The children indicated which emotion they thought was expressed by selecting a color photograph depicting one of the four emotions from a small group of photographs. Four- and 5-year-old American children were able to identify the emotion surprise, while American children 6 through 9 correctly identified all four emotions. Four- and 5-year-old Japanese children, on the other hand, were able to identify both surprise and sadness, while 6-year-olds correctly identified happiness, surprise, and sadness. Japanese children aged 7 through 9 were able to identify all four emotions. The results supported the concept of increasing decoding ability with increasing age across cultures, but also provided evidence for culture bound differences in growth of nonverbal emotional sensitivity.

Research on the development of nonverbal emotional sensitivity has been sparse until recently. Several authors have found that with increasing age, there is a general and gradual increase in the ability to identify emotional meanings communicated nonverbally across several channels, including the voice (Dimitrovsky, 1964) and the face (Buck, 1975; DePaulo & Rosenthal, 1979). Dimitrovsky (1964), for instance, found that children between the ages of 5 and 12 were able to correctly identify the negative emotional expressions of sadness and anger, followed by the positive emotions of happiness and love. Little follow up has been done, however, to replicate these findings.

The authors would like to express their gratitude to Robert Zajonc for his comments and suggestions on an earlier version of this manuscript; to Toshiko Tatsuno, Keiko Abe, Makiko Miyajima, and Linda Boerma for their assistance in the production of the stimuli used in the study; and to Akiko Enokido, who served as an experimenter.

Request for reprints may be sent to David Matsumoto, University of California, Psychology Clinic, 2205 Tolman Hall, Berkeley, CA 94720.

TABLE 1
Number of Subjects Tested at Each Age Level

Nationality	Age Groups				Total
	4 and 5	6	7	8 and 9	
American	9	14	11	13	47
Japanese	13	22	11	10	56

As an initial approach to investigating pancultural issues in children's ability to identify the emotional content of nonverbal vocal cues, the present study was carried out in the United States and in Japan. The Japanese culture was of particular interest, primarily because of the use of shame as a technique of social control, and its implications in children's identification of certain emotions such as sadness (Borke, 1973). Moreover, Benjamin (1977) has demonstrated that much information is communicated through the Japanese tone of voice, including emotional and situational determinants of conversation. As such the present study focused on the question of whether the ability to identify the emotional meanings of certain nonverbal vocal cues increases with age at the same rate across cultures.

METHOD

Subjects

The subjects were 110 children, ages 4 through 9. Of the 110 children, 60 were Japanese, while the remaining 50 were American. All of the children attended either a preschool or elementary school. Because of the small number of 4-, 5-, 8-, and 9-year-olds in both samples, the 4- and 5-year-olds were tested as one group, and their data were combined for statistical purposes, as were the 8- and 9-year-olds. Two of the 60 Japanese children were tested at the very beginning of the study as a check on the research methodology, as were 2 of the 50 American children. These data, therefore, are not included in the results. Also, two of the Japanese children and one American child declined to complete participation in the study, and consequently their data were not included either (see Table 1).

Materials

In studies examining the communication of emotion through paralinguistic cues, it becomes necessary to control for any effects words may have on the emotion judgments. Reducing the contribution of words to emotion communication enables the researcher to assess the independent

effects of the remaining paralinguistic cues such as tone of voice and intonation on such communication. Several methods now exist which enable the researcher to control for the effects of words, such as electronic speech filtering, splicing, the use of constant content, and content free speech (see Scherer & Ekman, 1982, for a comprehensive review). In the present study content free speech samples were used. Content free speech involves the production of voice samples consisting of numbers or alphabets. Each numeric or alphabet sample can be recorded so as to convey certain emotions, or have different emotional connotations.

In an attempt to replicate Dimitrovsky's (1964) findings on children's identification of emotions such as happiness, sadness, and anger, as well as Borke's (1973) pattern of results for sadness, these three emotions were included in this study. Moreover, cross cultural studies of facial expressions of emotion conducted after these earlier studies have supported the notion of the universal expression of other emotions, such as surprise, fear, and disgust (Ekman, 1982). As a preliminary extension of such work to the field of vocal communication of emotion, surprise was also included in this study.

In Japan 16 tapes were made with the use of content-free speech, each recording depicting one of the emotions of happiness, surprise, sadness, or anger. (Instead of the English alphabet, the Japanese syllabary was used.) Since recording the whole syllabary or alphabet would be too time-consuming and boring for the listener, only the first 15 syllables were recorded. Four recordings of each emotion were made by a female undergraduate who had previous experience in drama. Because it is often difficult to obtain adequate emotional stimuli from a "normal" population, we felt that the use of someone who was trained to express her emotions publicly and consistently would increase the chances of us obtaining an adequate sample of stimuli. All 16 tapes (four recordings per each emotion times four emotions) were subsequently reviewed by a group of six judges from the University of Tokyo Educational Psychology Department who, blind to the hypotheses of the study, selected the four recordings (one recording per each emotion) that were used in the remainder of the study. No recording that was selected had been mistakenly identified as another emotion, and the judges were unanimous in each of its choices.

In the United States the content-free speech tapes were also recorded by a female undergraduate student who had previous experience in drama. Again, four recordings, one of each emotion, were selected for the remainder of the study.

In dealing with young childrens' judgments of emotion it is necessary to utilize an adequate response instrument. Previous studies have incorporated a variety of methods, such as verbal reports or stick figure drawings (Borke, 1973; Dimitrovsky, 1964; Gates, 1923). In agreement with Borke

(1973) we thought that rendering the children's response repertoire to a behavioral rather than a verbal level would provide a more adequate response measure than simple verbal report. The empirical base now existing concerning pancultural facial expressions of emotion further offered us with a unique opportunity to utilize such a behavioral response measure.

Color photographs of faces expressing the four emotions were taken. The students who had done the speech recordings were also asked to pose in the four emotions. Rather than asking for a global emotional expression, the posers were asked to move specific muscle groups of the face, corresponding to the particular muscle patterns according to each emotion (Ekman & Friesen, 1975). Although it is difficult to elicit particular facial movements voluntarily at first, expressions corresponding quite well with Ekman and Friesen's (1975) notions of which areas of the face are innervated in certain emotions was achieved with some practice. Most students of facial behavior have had to work with many posers before obtaining adequate samples often because they did not have the theoretical or empirical foundation with which to guide their efforts as there now exists concerning the face (Ekman, 1982).

Four photographs of each emotion were taken. These photographs were also reviewed, and one of each emotion was selected for use in the remainder of the study. This procedure was repeated in both Japan and America. Again, no photograph that was selected had been mistakenly identified as another emotion, and the judges were unanimous in each of their choices. This extra screening procedure assured us of obtaining the purest sample available from our pool of possible stimuli. There were, therefore, altogether eight content free speech tapes and eight color photographs which were used in the testing.

Procedure

All of the subjects were tested individually. The child was first brought into a room, in which there were laid out on a table the four selected pictures and a tape recorder. The nationality of the pictures and tapes was kept congruent with the nationality of the subjects. The experimenter introduced herself, asked the age of the subject, and proceeded with the task.

The experimenter first directed the child's attention to the four photographs depicting happiness, sadness, surprise, and anger. The subject was then asked to identify the emotion depicted in each of the pictures. If a child had difficulty recognizing any of the faces, the experimenter would identify it for him or her. The experimenter would then proceed to describe the task, in somewhat the following manner.

Okay, now see this tape recorder here? Well, when I turn it on, this girl's (pointing to the pictures) voice will come out. When the tape finishes, I'd like you to pick out which girl you think said what's on the tape. The tapes are short, so you have to listen really well, now. Do you think you want to try?

The experimenter answered any questions the subject had, and then presented the child with the test tapes depicting each of the emotions. Following the presentation of each tape, the experimenter would again name the four photographs and ask the child to select the one that best exemplified the speaker of the passage. The picture the child selected was recorded, and the procedure was again repeated until each of the four tapes of the same nationality were played. The order of the tape presentation for each subject was kept random and counterbalanced throughout the entire experiment so that each tape was heard in each of the four possible positions an equal number of times. After all four tapes of the same nationality were played, the child was then told that that was all, and was thanked for participating in the study.

RESULTS

The percents of correct scores for each of the four emotions across each of the age groups were computed for both the American and Japanese samples, and are presented graphically in Figures 1 and 2. All 47

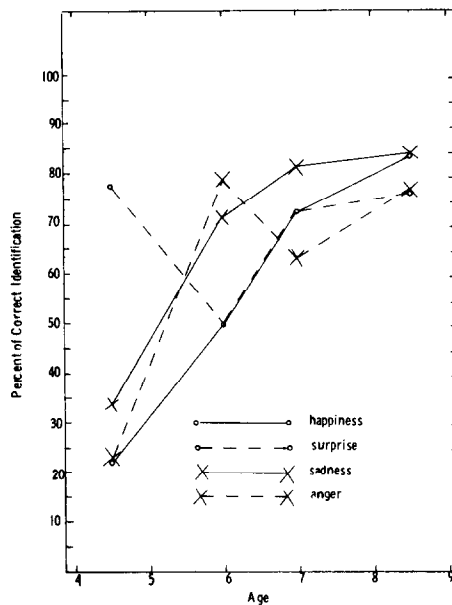


FIGURE 1. Percent of correct identification per each emotion per each age group, American children.

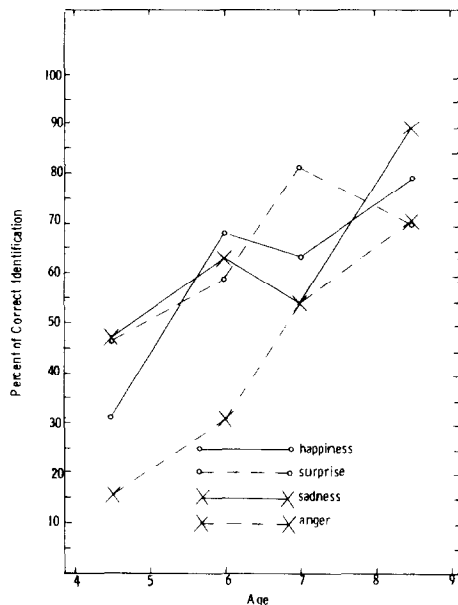


FIGURE 2. Percent of correct identification per each emotion per each age group, Japanese children.

American children identified each of the four emotions upon hearing the test tape beyond levels expected by chance alone ($\chi^2(3) = 124.13, p < .001$). Likewise, all 56 Japanese children identified each of the four emotions conveyed in the test tape at a significant level ($\chi^2(3) = 96.14, p < .001$). When we examine these data at each age level, however, we find some evidence for differential degrees of emotional sensitivity. The 4- and 5-year old American children, could only identify surprise at a significant level ($\chi^2(1) = 10.03, p < .01$), but could not identify the other three emotions ($\chi^2(2) = .306, ns$). The 4- and 5-year-old Japanese children, on the other hand, were able to correctly identify not only surprise, but also sadness ($\chi^2(1) = 4.65, p < .05$).

The 6-year-old American children correctly identified all four emotions from the nonverbal vocal stimuli ($\chi^2(3) = 35.14, p < .01$). The 6-year-old Japanese subjects, however, were able to identify only happiness, surprise, and sadness ($\chi^2(2) = 39.78, p < .01$).

The 7-year-old Americans ($\chi^2(3) = 40.81, p < .001$), as well as the 8- and 9-year-olds ($\chi^2(3) = 65.00, p < .001$) were able to correctly identify all four emotions. Likewise, the 7-year-old Japanese children ($\chi^2(3) = 28.45, p < .001$), as well as the 8- and 9-year-olds ($\chi^2(3) = 45.20, p < .001$), correctly identified all four emotions at levels beyond that expected by chance.

The Japanese and American data were compared to examine the presence of cultural differences in sensitivity. This was done across each of the age groups studied, as well as across each of the four emotions. The pattern of results reported by the American and Japanese samples were significantly different only in the 4- and 5-year-old age group ($\chi^2(1) = 7.21, p < .01$). All other age by culture interactions were not significant. Examination of each of the four emotions across cultures, however, revealed that American children consistently identified anger more often than Japanese children ($\chi^2(1) = 6.16, p < .025$). Other emotion by culture interactions were not significant.

DISCUSSION

Age Differences

The results of this study indicate that, between the ages of 4 and 9, there is a gradual and generally steady increase in the ability to identify the emotional meaning of nonverbal vocal expressions of the four emotions. This general finding is consistent with similar earlier findings (De Paulo & Rosenthal, 1979; Dimitrovsky, 1964; Gates, 1923). It is apparent, then, that the childhood years of 4 through 9 represent a period of considerable growth in sensitivity to emotional communication, with accuracy levels rising from about 30% to 80%. The trend of increasing decoding ability with age continues throughout each of the successive ages, as all four of the emotions were correctly identified with increasing accuracy rates.

Cultural Differences

The finding that American children 6 years of age through 9 most frequently identified sadness and anger is consistent with the findings of Dimitrovsky (1964). Dimitrovsky, however, did not test children younger than 5; in our 4- and 5-year-old American sample, surprise was identified at a much higher rate than sadness and anger. This discrepancy may be attributable to several factors. First of all, the sample size of the 4- and 5-year-old group was small ($n = 9$), which makes generalizations from the data less reliable than if a larger sample size was tested. But because of the large percentage of subjects who did correctly identify surprise, as well as the fact that the Japanese 4- and 5-year-olds were able to identify surprise, it may very well be that of the four emotions tested, surprise rather than sadness or anger is most easily identified below this age.

When we consider the Japanese children's accuracy rates for anger, we find that they show an increasing accuracy in their identification of anger from age 4 through 9. The overall rate, however, is low; it is lower

than their rates of accuracy for the identification of the other three emotions, as well as lower than American children's accuracies of identification for all emotions. Low Japanese accuracy rates on anger could represent a true cultural difference. The difference, however, may not lie in the Japanese subjects' accuracy of judgment of anger but rather in their response behavior. That is, even if the Japanese children identified anger through both the nonverbal vocal cues and the photographs as well as American children, they may not as easily give the actual response indicating that they perceived anger. This process could be related to the effects of the early Japanese socialization process, where children are socialized to withhold such emotions as anger (Benedict, 1946). That is, they do not as readily as their American counterparts communicate feelings of anger, as they "bottle it up." This same effect may be interacting in tests presented in a study such as this. That is, even if the correct identification of anger can take place, there may be greater inhibition for a Japanese child to convey that identification to an experimenter (similar to the mechanism underlying display rules and facial expressions; see Ekman, 1972, 1977). Perhaps the recognition of even a vocal quality not particular to anger but socialized to produce such an effect would cue such response inhibition. The effects of shame within the socialization process, or any effects produced by the differential socialization activities between the societies, therefore, may produce cultural variations not only in the correct identification of emotions (such as sadness, as pointed out earlier) but can also provide for variation in misidentification, or more precisely, response (or nonresponse) patterns concerning emotions such as anger.

While many of the effects evidenced in both samples were strong, the interpretation of these effects must be made with caution concerning the test materials, the particular response repertoire available to the younger subjects, and the sample size. It may very well be that the utilization of more sensitive types of stimuli produced by a wider range of people will result in more pronounced effects. In any case the results of the present study may reflect differences in individual emotional development and in the socialization practices of the two countries, and nevertheless provides a springboard for future research of this nature with different populations. Other data suggest that children even younger than 4 are able to correctly identify nonverbal vocal expressions of emotion (Tatsuno, 1981). Testing the development of emotional sensitivity of younger children should have important implications to understanding not only socialization practices, but also the development of individual personality differences.

REFERENCES

- BENEDICT, R. *The chrysanthemum and the sword: Patterns of Japanese culture*. New York: Houghton Mifflin & Company, 1946.

- BENJAMIN, G. Tone of voice in Japanese conversation. *Language and Society*, 1977, 6, 1-13.
- BORKE, H. The development of empathy in Chinese and American children between three and six years of age. *Developmental Psychology*, 1973, 9, 102-108.
- BUCK, R. Nonverbal communication of affect in children. *Journal of Personality and Social Psychology*, 1975, 31, 644-653.
- DePAULO, B., & ROSENTHAL, R. Age changes in nonverbal decoding skills: Evidences for increasing differentiation. *Merrill-Palmer Quarterly of Behavior and Development*, 1979, 25, 145-150.
- DIMITROVSKY, L. The ability to identify the emotional meaning of vocal expressions at successive age levels. In J.R. Davitz (Ed.), *The communication of emotional meaning*. New York: McGraw-Hill, 1964.
- EKMAN, P. Universal and cultural differences in facial expressions of emotion. In J.K. Cole (Ed.), *Nebraska Symposium on Motivation (Vol. 20)*. Lincoln: University of Nebraska Press, 1972.
- EKMAN, P. Biological and cultural contributions to body and facial movements. In J. Blacking (Ed.), *The anthropology of the body*. London: Academic Press, 1977.
- EKMAN, P. (Ed.), *Emotion in the human face*. New York: Cambridge University Press, 1982.
- EKMAN, P., & FRIESEN, W. *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall, 1975.
- GATES, G. An experimental study of the growth of facial perception. *Journal of Educational Psychology*, 1923, 14, 449-461.
- SCHERER, K., & EKMAN, P. (Eds.), *Handbook of methods in nonverbal behavior research*. New York: Cambridge University Press, 1982.
- TATSUNO, T. Kanjo Nenchi no Hattatsutekina Kenkyu (Developmental Research on the Perception of Emotions). *Kyoshoku, Shaki Kyoiku Shugi, Gakugeiin Katei Nempo*. Tokyo, Japan: Meiji University Press, 1981.

ABSTRACT TRANSLATIONS

Dans une étude sur le développement du jugement à partir de perception auditive, des bandes magnétiques, sur lesquelles sont enregistrées des discours exprimant une des quatre émotions suivantes bonheur, surprise, tristesse et colère, ont été présentés à 47 enfants américains et 56 japonais âgés de 4 à 9 ans. Les enfants indiquaient celle des émotions qu'ils pensaient être exprimées en désignant une photo en couleur, parmi plusieurs, dépeignant une des 4 émotions. Les enfants américains âgés de 4 et 5 ans ont été capables d'identifier l'émotion de surprise tandis que ceux de 6 et 9 ans ont correctement identifiés les 4 émotions. Par ailleurs les japonais de 4 et 5 ans, ont reconnu la tristesse et la surprise alors que ceux de 6 ans ont reconnu le bonheur, la surprise et la tristesse. Les japonais de 7 et 9 ans ont également identifiés les 4 émotions. Ces résultats confirment l'idée que la faculté de interprétation augmente avec l'âge quelque soit la culture, et aussi ces résultats fournissent une preuve de la différence de développement avec la culture de la sensibilité et des erreurs négligées.

En un estudio exploratorio de la tendencia del desarrollo de juicios emocionales, a partir de señales orales no verbales, es contenido en estas cintas magnetofónicas. Estas cintas fueron diseñadas para que el niño exprese una de las cuatro emociones de alegría, sorpresa, tristeza o enojo que son expresadas por cuarenta y siete niños Americanos y cincuenta y seis niños Japoneses, entre las edades de cuatro a nueve años. Los niños indicaban la emoción que ellos creían que era expresada en una selección de un pequeño grupo de fotografías de color en las cuales se expresaban dichas emociones. Los niños Americanos con la edad correspondiente de cuatro a cinco años tuvieron la habilidad de identificar las emociones de sorpresa. Niños Americanos con la edad de seis a nueve años identificaron correctamente las cuatro emociones. Los niños Japoneses con la edad correspondiente de cuatro a cinco años, por el contrario, fueron capaces de identificar las emociones de sorpresa y tristeza, mientras tanto los de seis años de edad identificaron correctamente las emociones de alegría, sorpresa y tristeza. Niños Japoneses con la edad de siete años también fueron capaces de identificar las cuatro emociones. Este resultado apoya el concepto del incremento de la habilidad de captar a medida que la edad de los niños aumenta en ambas culturas, pero también provee un resultado de la diferencia en el crecimiento de la sensibilidad de lo no verbal y los patrones de error.