Journal of Experimental Child Psychology 102 (2009) 78–95

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Children's perception and understanding of (dis)similarities among dynamic bodily/facial expressions of happiness, pleasure, anger, and irritation

Sandrine Vieillard *, Michèle Guidetti

Université Toulouse II–Le Mirail, Unité de Recherche Interdisciplinaire OCTOGONE, Laboratoire Cognition, Communication, et Développement (ECCD), 31058 Toulouse Cedex 9, France

A R T I C L E I N F O

Article history:
Received 31 October 2007
Revised 22 April 2008
Available online 12 June 2008

Keywords:
Emotions
Dynamic expressions
Free classification
Free labeling
Multidimensional scaling
Cluster analysis

A B S T R A C T

The current study examined the abilities of children (6 and 8 years of age) and adults to freely categorize and label dynamic bodily/facial expressions designed to portray happiness, pleasure, anger, irritation, and neutrality and controlled for their level of valence, arousal, intensity, and authenticity. Multidimensional scaling and cluster analyses showed that children ($n = 52$) and adults ($n = 33$) structured expressions in systematic and broadly similar ways. Between 6 and 8 years of age, there was a quantitative, but not a qualitative, improvement in labeling. When exposed to rich and dynamic emotional cues, children as young as 6 years can successfully perceive differences between close expressions (e.g., happiness, pleasure), and can categorize them with clear boundaries between them, with the exception of irritation, which had fuzzier borders. Children's classifications were not reliant on lexical semantic abilities and were consistent with a model of emotion categories based on their degree of valence and arousal.

Introduction

One of the critical issues in emotional development is the decoding of emotions to assess the social environment quickly and modify behavior accordingly, thereby ensuring successful social communication. The human face has a special status because it can be processed faster and more efficiently than any other class of objects (Farah, Wilson, Drain, & Tanaka, 1998; Kanwisher, 2000). This status

* Corresponding author. Fax: +33 5 61 50 39 19.
E-mail address: sandrine.vieillard@univ-tlse2.fr (S. Vieillard).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2008.04.005

endows the human face with greater significance in social communication. Thus, it is not surprising that most early studies of emotional development focused on facial expressions. This widespread interest overshadowed research on emotional body movements, which have remained a neglected area ever since. Although it is well known that nonverbal emotional communication depends on both bodily and facial expressions, up to now these two media have been considered separately for research purposes.

Previous research on children's ability to decode emotion in expressive body movements had suggested that this ability did not emerge until 8 years of age (Custrini & Feldman, 1989; Van Meel, Verburgh, & DeMeijer, 1993). More recently, however, research has shown that children as young as 4 years can nonverbally identify discrete emotional meaning in expressive body movements at above-chance levels and that children as young as 5 years display an increased ability to identify these emotional meanings (Boone & Cunningham, 1998). These data demonstrate the success of kinesic displays in communicating emotional meaning even to young children.

Regarding the development of facial emotion processing, research has shown that the ability to use specific facial muscle configurations to discriminate facial expressions is present at birth (e.g., Field, Woodson, Greenberg, & Cohen, 1982) and that emotion perception skills develop during the first year of life. Although 4- to 6-month-olds are able to discriminate facial expressions of anger, fear, surprise, and sadness (e.g., Serrano, Iglesias, & Loeches, 1992; Walker-Andrews, 1998), the ability to abstract facial expressions from different people does not emerge until approximately 7 months of age (e.g., Caron, Caron, & Myers, 1982). The ability to perceive differences in facial expressions develops early in human ontogeny, but the ability to recognize emotion per se follows a slower developmental course, lasting from 3 to 10 years of age (e.g., Camras & Allison, 1985; De Sonneville et al., 2002; Durand, Gallay, Seigneuric, Robichon, & Baudouin, 2007; Gosselin, 1995, 2005; Herba, Landau, Russell, Ecker, & Phillips, 2006; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000). Recently, however, promising findings have provided evidence that 4-month-olds may have a rudimentary understanding of expressions of sadness, anger, and fear (Montague & Walker-Andrews, 2001), suggesting that the ability to recognize emotion per se may emerge before 3 years of age.

Although these studies have used different methodologies and produced conflicting findings on the developmental course of facial emotion recognition, they nevertheless indicate that emotion recognition processing does not emerge as a complete package and that the developmental pattern is not uniform across all emotions. Typically, happiness is the earliest and most accurately recognized emotion, followed by sadness, with less clear patterns for anger, fear, surprise, and disgust (e.g., Boyatzis, Chazan, & Ting, 1993; Camras & Allison, 1985; Durand et al., 2007; Gosselin, 1995, 2005; Vicari et al., 2000). There are several explanations as to why happiness is recognized so early and so accurately. It has been suggested that the lower part of the face—a smile—is sufficient to identify happiness, whereas other emotions—such as fear, anger, and sadness—require a combination of both the upper and the lower parts of the face (Vicari et al., 2000). Other authors have pointed out that there are no other positive emotions in the set of basic emotions with which happiness must compete (e.g., De Sonneville et al., 2002) and that happy faces are those most frequently seen by most children and, thus, are the most familiar and readily processed faces (Batty & Taylor, 2006).

A common feature of these studies has been the investigation of emotional responses to facial emotions using basic expressions—happiness, sadness, anger, fear, disgust, and surprise. This choice of basic emotions is based on the traditional discrete approach, according to which these emotions correspond to a limited number of innate and universal emotion categories from which all other emotions can be derived (Ekman, 1982; Izard, 1997). In this approach, happiness, fear, sadness, anger, and disgust are considered as discrete emotions in that they are assumed to be unique emotional states that stem from distinct causes (e.g., Izard, 1997). These emotions are viewed as triggering basic and distinct adaptive behaviors. Support for this approach can be found in neuropsychological and functional brain imaging studies. For example, the recognition of facial expressions of fear may be linked to specific neural substrates such as the amygdala (Adolphs, Tranel, Damasio, & Damasio, 1994; Calder et al., 1996; Morris et al., 1996; Phillips et al., 1998), whereas the recognition of facial expressions of disgust may be related to the basal ganglia and the anterior insula (Gray, Young, Barker, Curtis, & Gibson, 1997; Phillips et al., 1998; Sprengelmeyer et al., 1996). However, the discrete emotions approach does not account for the possibility of strong degrees of co-occurrence between discrete emotional states (mixed emotion). Furthermore, it appears that the ability to recognize discrete emotions may vary as a function of the participants' origins (Russell, 1994), thereby challenging the universality view.

Another traditional approach stipulates that emotions may be characterized along a smaller number of continuous psychological dimensions. One of the most familiar models is Russell's (1980) circumplex model of affect, which specifies that there are two main affective dimensions reflecting degrees of valence and arousal. Numerous previous studies have established that these two main dimensions of valence and arousal may structure facial expressions (Abelson & Sermat, 1962; Osgood, 1966), voice perception (Green & Cliff, 1975), affect words (Russell, 1980), and music (Bigand, Vieillard, Madurell, Marozeau, & Dacquet, 2005; Vieillard et al., in press). Furthermore, developmental studies using multidimensional scaling (MDS) analysis have demonstrated that the emotional responses of both young children (2–5 years of age) and older ones (11–14 years of age) to facial expressions are structured in a two-dimensional space of valence and arousal (Bullock & Russell, 1984, 1985; Russell & Ridgeway, 1983), indicating that these two dimensions reflect basic processes underlying their interpretation of facial expressions. These empirical findings suggest that facial expressions may initially be perceived in terms of pleasure and arousal rather than in terms of clear-cut categories of prototypical emotions (Bullock & Russell, 1984, 1985). To date, the question of whether the discrete approach or the dimensional approach to emotion theory is most appropriate remains unanswered. In the current study, we attempted to combine these two approaches using both categorical (discrete) and dimensional methodologies to investigate emotional perception in children.

In an effort to understand how children interpret emotional messages conveyed by facial expressions (for a review, see Gosselin, 2005), many researchers have used photographs, drawings, or stories, asking children to perform a *forced-choice* labeling task (e.g., Bullock & Russell, 1985; Camras & Allison, 1985; Gosselin, 1995). It has been found that children as young as 2 years can already perform correctly at above-chance levels for many expressions (e.g., fear, happiness, sadness). By 5 years of age, and depending on task demands (target/distractor ratio), their performance can be on a par with that of adults for all basic emotions. Using a *free* labeling task, Widen and Russell (2003) showed that the developmental course of verbal emotional categorization runs more slowly. These authors found that 2-year-olds failed to label facial expressions correctly and that some expressions (e.g., disgust) remained problematic until approximately 5 years of age. This suggests that children's abilities to interpret facial expressions continue to develop beyond the 5th year and may follow a slow pace. The developmental course of label emergence indicates that "happy" comes first, followed by "angry" and "sad" (Widen & Russell, 2003). To date, there has not been any research on older children's abilities to freely assign correct labels as a function of dynamic and finely tuned bodily/facial expressions.

Another striking point is that most developmental studies of facial expressions have employed static stimuli (i.e., photographs). This is potentially problematic in that still faces may have less ecological validity. Some empirical findings suggest that dynamic displays improve the recognition and discrimination of emotions compared with static ones (Lemay, Kirouack, & Lacouture, 1995; Werhle, Kaiser, Schmidt, & Scherer, 2000) and that they involve different neural pathways (Kilts, Egan, Gideon, Ely, & Hoffman, 2003). Accordingly, it may be that findings based on static emotional material underestimate children's emotional perceptual skills. More significant, in their daily lives, laypeople are used to seeing dynamic facial expressions, not static ones. Using a naturalistic context of "Peekaboo," Montague and Walker-Andrews (2001) reported that infants as young as 4 months were able to discriminate and respond in meaningful ways to dynamic expressions of sadness, anger, and fear. These findings not only showed that young infants can perceive the differences among sadness, anger, and fear but also indicated that they may respond differentially to expressions, suggesting a rudimentary understanding of the expression per se. However, because these authors incorporated both visual and auditory modality in their Peekaboo paradigm, it is impossible to disentangle the respective contributions of these modalities to abilities to decode expressions. Regarding the visual modality, emotion expression involves the synchronization of facial movements and body postures, both of which are usually displayed to differing, but rarely extreme, degrees of intensity. Because no research has directly and jointly addressed these issues to date, healthy children's ability to decode and interpret subtle dynamic displays of adjacent bodily/facial expressions (e.g., happiness, pleasure) remains unknown.

Current study

The main purpose of the current study was to examine how children use valence and arousal cues in dynamic bodily/facial expressions to discriminate the sort of human emotions they see around them in real life.[1] A second important objective was to test how children perceive expressions pertaining to the same family but conveying different degrees of arousal. A further aim was to examine how emotional classification and labeling develop between 6 and 8 years of age, a period during which improvements in emotion processing have been observed (Durand et al., 2007; Gosselin, 1995; Gosselin & Larocque, 2000; Herba et al., 2006; Vicari et al., 2000).

Expressions of happiness, pleasure, anger, irritation, and neutrality were selected for a number of reasons. First, the literature on emotional development has recurrently reported that positive basic expressions of happiness are recognized better than negative ones. The choice of happiness and pleasure as two related emotions, differentiated only by their degree of arousal, was an attempt to challenge the traditional method of using a single, undifferentiated positive state (e.g., happiness) to test emotion processing. Second, happiness and anger were chosen as two opposite emotions. One aim was to test whether a happy expression would be identified better than an angry one in an emotional context opposing two positive expressions (i.e., happy and pleased) and two negative ones (i.e., angry and irritated). Finally, the use of expressions pertaining to the same family of emotions, but varying in their degree of arousal (e.g., anger, irritation), was a means of crossing valence and arousal factors to disentangle their respective influence on children's perception and understanding of (dis)similarities among dynamic bodily/facial expressions. Neutral expressions, which corresponded to nonemotional portrayals, were included to test how children process inexpressive body/face configurations presented alongside expressive ones. In the emotion literature, there is a consensus that happiness and anger are discrete emotions, whereas pleasure and irritation are defined more in terms of valence and arousal. Consequently, pleasure and irritation are not clear emotion categories. With this in mind, we decided to use "expression category" to designate each of the four expressions featured in the current study.

A free classification task and MDS and clustering methods of analysis were chosen to test whether the boundaries among the expression categories were fuzzy or clear-cut. This methodology goes beyond the traditional one of assessing the accuracy of a child's performance by measuring it against an adult's standards, thereby possibly underestimating the child's ability to perceive (dis)similarities among expressions. A free classification task was also used because of its relative simplicity and the ecological relevance of spontaneous categorization. This allows participants to choose their own similarity criteria to create different classes without needing to verbalize; participants create classes based directly on the perceived stimulus cues without needing to label the behavior being displayed. During a second phase, the classification task was followed by a labeling task designed to assess lexical semantic abilities. A free labeling task was preferred to a forced-choice one because the latter might have yielded artifactual results due to the fact that it does not allow participants to spontaneously choose labels other than the ones proposed by the experimenter. A free labeling task should also provide some cues as to how children verbally categorize expressions that are similar in valence but differ in arousal.

The adults' performances on the free categorization and verbalization tasks were used as a benchmark for comparisons with the children. Given that the portrayals that had been rated beforehand in the pilot study were intended to communicate expressions of happiness, pleasure, anger, irritation, and neutrality, the adults were expected to create five groupings on average and to use emotional labels that properly designated each intended expression. Regarding the children's performances, because emotion processing has generally been found to improve with age, we expected the free classification to be performed better by the 8-year-olds than by the 6-year-olds. According to the dimensional conception of emotions, the children would discriminate the expressions on the basis of their valence and arousal values. More specifically, because 6-year-olds' emotion processing abili-

[1] The emotional material used in this study did not take the form of a three-dimensional display as in real life; rather, it took a two-dimensional dynamic display of body movements synchronized with facial expressions.

ties were found to be close to those of adults for happiness, it was expected that younger children would first distinguish "happy" expressions from "unhappy" ones and then discriminate stimuli as a function of their valence. Accordingly, children as young as 6 years should create at least three different classes: pleasant, unpleasant, and neutral expressions. Because 6- and 8-year-olds have been found to correctly discriminate negative emotions with differing degrees of arousal (e.g., anger, sadness), they were expected to use the arousal dimension as a means of discriminating between expressions such as anger and irritation. As for lexical semantic abilities, previous findings had indicated that, in free labeling tasks, the production of correct labels slowly improves with age. Thus, no major improvements in the production of correct labels were expected. However, we hypothesized that if younger children focused on valence first, their labeling would reflect this and would be based on pleasant–unpleasant terms.

Method

Participants

Three different age groups participated in the experiment. The 85 total participants consisted of 28 6-year-olds (14 girls and 14 boys, mean age = 6 years 2 months, range = 5 years 7 months to 6 years 6 months), 24 8-year-olds (14 girls and 10 boys, mean age = 8 years 1 month, range = 7 years 7 months to 8 years 6 months), and 33 adults (16 women and 17 men, mean age = 29 years 9 months, range = 21 years 11 months to 41 years 8 months). The children were from predominantly middle-class families and shared homogeneous ethnic characteristics (47 Caucasian, 2 Afro-French, and 3 Asian French). All were recruited from public elementary schools in Toulouse, France. Parents were sent a letter describing the study along with a consent form. Only children with informed parental consent participated in the study. Adult participants, all of whom volunteered to take part in the study, were Caucasian French psychology or chemistry graduate students or researchers.

Apparatus

The children were tested in a quiet room at their school, and the adults were tested in a university office. The stimuli were presented and classified using PowerPoint software on a personal computer. Each participant performed the task individually and was seated in front of a 15.4-inch monitor controlled by the computer at a mean distance of 60 cm.

Materials

A total of 40 dynamic visual stimuli were taken from the Geneva Multimodal Emotion Portrayals corpus (GEMEP) (Bänziger, Pirker, & Scherer, 2006) containing 143 silent videotapes, each lasting 1 s. The videotapes showed the facial and bodily expressions of male and female actors—two females (one young and one middle-aged) and two males (one young and one middle-aged)—who had been requested to portray each of the four intended expressions of happiness, pleasure, anger, and irritation. Additional silent video sequences, in which actors did not express any emotion at all, were selected from GEMEP to obtain a neutral condition. In these sequences, no specific instructions were given to the actors; they simply waited and prepared for a new emotion. We selected sequences that were as void of emotion as possible.

Stimulus selection procedure

In a pilot study described in Appendix A, all stimuli were rated on a 10-point scale for their level of emotional clarity, valence, arousal, intensity, and authenticity. The rationale for using arousal and intensity ratings was to test whether these dimensions would reflect different facets of the same expression. Distinguishing arousal from intensity seemed particularly convenient for stimuli depicting emotional body movements that can be more or less vigorous. For instance, an expression of irritation can be judged as low arousal because the body movements lack energy and yet can be judged to be highly intense for the reason that it is perceived as communicating a deep emotional experience.

Because arousal and valence are known to be the two main dimensions that govern emotional responses, a mixed repeated-measures analysis of variance (ANOVA) was performed on these ratings for each videotape with expression category as a between-items factor and arousal and valence as within-items factors. There was a significant interaction between expression categories and ratings, $F(4, 35) = 135.92$, $p < .001$, $\eta^2 = 1$. Post hoc Bonferroni tests showed that happiness ($M = 6.89$, $SE = 0.32$) and pleasure ($M = 1.86$, $SE = 0.32$), as well as anger ($M = 7.84$, $SE = 0.32$) and irritation ($M = 3.61$, $SE = 0.32$), all differed significantly from each other along the arousal dimension ($p < .001$). Furthermore, positive emotions such as happiness ($M = 7.69$, $SE = 0.16$) and pleasure ($M = 6.55$, $SE = 0.16$) differed significantly ($p < .001$) from negative expressions such as anger ($M = 0.95$, $SE = 0.16$) and irritation ($M = 2.30$, $SE = 0.16$) along the valence dimension. Neutral stimuli differed significantly from all of the others on both the arousal ($M = 0.61$, $SE = 0.32$) and valence ($M = 3.58$, $SE = 0.16$) scales ($p < .05$).

The mean ratings of valence and arousal for each of the 40 selected videotapes are plotted in Fig. 1. As can be seen, the participants systematically rated the expressions of happiness and pleasure as pleasant and rated the expressions of anger and irritation as unpleasant. With regard to the arousal dimension, expressions of pleasure and irritation received a relatively low rating, whereas expressions of anger and happiness received a high rating. Neutral stimuli were rated at an intermediate level on the valence axis and at a low level on the arousal one. This distribution is consistent with the motivational model developed by Lang, Bradley, and Cuthbert (1999) in which stimuli are plotted in a boomerang shape, with the two arms reaching toward the high-arousal quadrants. In this model, the arms correspond to the appetitive and defensive systems that are thought to reflect the motivational foundation of affective judgments.

Two separate one-way ANOVAs indicated that expression categories affected the judgments of authenticity, $F(4, 35) = 3.864$, $p < .05$, $\eta^2 = .85$, and intensity, $F(4, 35) = 69.482$, $p < .001$, $\eta^2 = 1$. For authenticity ratings, post hoc Bonferroni tests showed that the mean score for neutral expressions ($M = 4.63$, $SE = 0.31$) differed significantly from that for anger ($M = 6.29$, $SE = 0.31$) but not from that

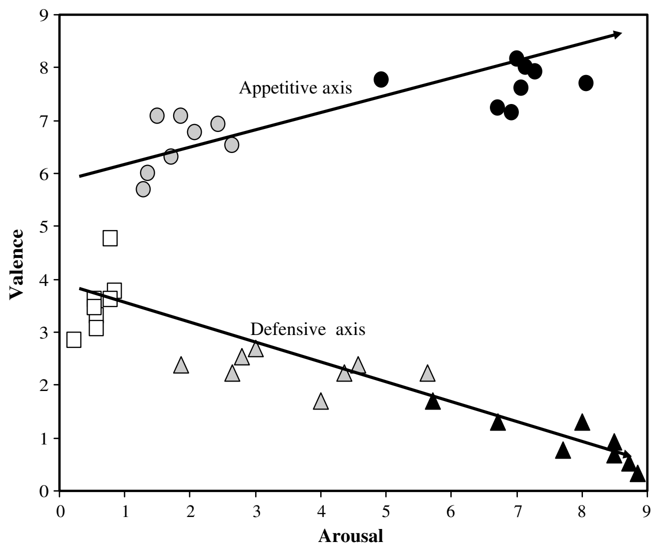

Fig. 1. Plots of 40 selected videotapes on the basis of their mean scores of valence (y axis) and arousal (x axis). Each point represents the ratings for portrayals of happiness (black circle), pleasure (gray circle), anger (black triangle), irritation (gray triangle), and neutrality (white square). The upper arrow represents the motivational system of appetite, and the lower arrow represents the motivational system of defense (Lang et al., 1999).

for happiness ($M = 5.97$, $SE = 0.31$), pleasure ($M = 5.01$, $SE = 0.31$), or irritation ($M = 5.23$, $SE = 0.31$). For intensity ratings, post hoc Bonferroni tests showed significant differences among the expression categories except between pleasure and happiness and between irritation and anger.

To test whether the intensity and arousal scales assessed the same facet of the stimuli, a repeated-measures ANOVA, with expression categories as a within-items factor and arousal and intensity ratings as between-items factors, was performed on the mean score for each stimulus. There was a significant interaction between expression categories and ratings, $F(4, 35) = 17.01$, $p < .001$, $\eta^2 = 1$. Intensity ratings for irritation ($M = 4.58$, $SE = 0.30$), pleasure ($M = 4.19$, $SE = 0.30$), and neutrality ($M = 1.61$, $SE = 0.30$) were higher than the respective arousal ratings (irritation: $M = 3.61$, $SE = 0.32$; pleasure: $M = 1.86$, $SE = 0.32$; neutrality: $M = 0.53$, $SE = 0.32$), indicating that the arousal and intensity scales did not assess exactly the same stimulus features. The mean ratings for the clarity, valence, arousal, intensity, and authenticity of each portrayal are presented in Appendix A.

Procedure

The 40 selected stimuli were divided into two sets of 20 videotapes so that each set included portrayals of all five expression categories (happiness, pleasure, anger, irritation, and neutrality) by all four actors (one young female, one middle-aged female, one young male, and one middle-aged male). All participants were then presented with a random visual pattern of 20 videotapes displayed on a PowerPoint interface with a slide size of 1.575×0.787 inches.

The adults were instructed to look for videotapes portraying people who displayed a similar emotional experience and then to group them together. The adults needed to click on a stimulus of their choice to increase the size of the display and play the tape. After it was over, the display recovered its original size and the adults continued to watch the videotapes with a view to grouping together the expressions they perceived as being emotionally similar by dragging and dropping them into the same location. Adults could put any number of videotaped pictures into a group and could make as many groups as they liked. Once they were satisfied with their classifications, they were asked to perform a free labeling task, specifying verbally and with a single term the emotional similarities among the videotaped pictures within each group. They were asked to enter labels next to each group aid of the keyboard.

Because not all of the children were equally at ease with the use of a mouse on a computer, we used a testing method whereby the experimenter assisted them. The procedure was quite similar to the one for adults except that the computer was handled by the experimenter as instructed by the children. The children were told the following: “Look at the screen. Show me the men/women who feel the same way and that you want to put together. Don’t look for the same person but for people who are feeling the same.” The children were free to choose the videos that they wanted to watch so as to ensure random stimulus selection across participants. Once a child had made his or her choice, the experimenter clicked on the videotaped picture and systematically compared the new stimulus with those already seen, asking, “Do you think that these two people feel the same way?” If the child answered “yes,” the experimenter asked, “Should we put them together?” and waited for the child’s confirmation. If the child answered “no,” the experimenter queried “We don’t put them together?” and waited for the child’s confirmation. Once the classification task was over, the children were first asked to check their groupings and to say what the people in each group felt. To minimize the direct influence of the experimenter’s behavior on their responses, the experimenter sat to one side of the children but slightly behind so as to be out of sight, spoke in a controlled tone throughout the session, and kept verbal interactions to a minimum; that is, the experimenter entered the children’s verbal labels without providing any feedback about the responses. Each testing session lasted between 30 and 40 min. All participants were debriefed at the end of the experiment.

Results

Two participants (one 6-year-old and one adult) were excluded from the analysis because of a technical problem.

Number of groupings

The number of groupings was computed to test the general assumption that emotion classification would be enhanced with age. The 6-year-olds produced a mean number of 3.9 groupings ($SE = 0.19$), the 8-year-olds produced 4.8 groupings ($SE = 0.21$), and the adults produced 5.6 groupings ($SE = 0.18$). A two-way ANOVA, with the number of groupings as a dependent factor and age and set of stimuli as independent factors, showed a significant effect of age, $F(2, 77) = 22.38$, $p < .001$, $\eta^2 = .99$. There was a significant linear trend, $F(1, 77) = 1886.63$, $p < .001$, showing that the adults created a greater number of groupings than the 8-year-olds, who in turn created a greater number of groupings than the 6-year-olds. The mean number of groupings increased with age by approximately 12% between the 6- and 8-year-olds and by approximately 10% between the 8-year-olds and the adults. No significant effect of the set of stimuli or interaction between the age and set of stimuli factors was found.

Multidimensional scaling

We used an MDS analysis to examine the dimensions that guide children's categorization of emotions with increasing age. As outlined above, younger children were expected to discriminate expressions chiefly along the valence dimension, whereas older children were expected to do so along both the valence and arousal dimensions, thereby allowing them to distinguish among the five intended expressions.

For each set of stimuli, the groupings made by participants were converted into a 20×20 co-occurrence matrix. Each cell of the matrix indicated the average number of times for which two videotaped pictures were grouped together. The mean of the co-occurrence matrix was subtracted from 1 to produce a dissimilarity matrix. Thus, in each cell, the smaller the value, the more similar were the two stimuli. Given that groupings were based on two different sets of videotaped pictures, correlation analyses were performed to ascertain whether the responses were consistent across set conditions. Results indicated that the matrices obtained for the two sets were highly correlated for the 6-year-olds, $r(398) = .91$, $p < .001$, the 8-year-olds, $r(398) = .91$, $p < .001$, and the adults, $r(398) = .92$, $p < .001$. Given this degree of consistency, the two matrices were merged for each group of participants; the dissimilarity values obtained for each of the two sets were averaged for each expression category, resulting in a final 20×20 matrix for each age group. The correlation analyses showed that the matrix obtained for the 6-year-olds was highly and positively correlated with that for the 8-year-olds, $r(398) = .97$, $p < .001$, and that for the adults, $r(398) = .89$, $p < .001$, which in turn was highly correlated with that for the 8-year-olds, $r(398) = .99$, $p < .001$.

MDS analyses were applied to examine relations between the stimuli as a function of the expression categories for each age group. Matrices obtained for the 6-year-olds, 8-year-olds, and adults were processed separately using the PROXSCAL MDS algorithm in SPSS software (Kinnear & Gray, 1999). This yielded three two-dimensional solutions in a common geometrical space (using Torgerson's classic metric analysis), and these were rotated to optimal congruence. The badness-of-fit measures for the MDS solutions, represented by the normalized stress value, reached .01, .02, and .04 for the 6-year-olds, 8-year-olds, and adults, respectively. This means that more than 95% (99, 98, and 96% for the 6-year-olds, 8-year-olds, and adults, respectively) of the distances between the points located in the three geometrical solutions matched the rank order of the dissimilarities in their related proximity matrix. The percentages of variance explained by the MDS solutions were 98, 97, and 95% for the 6-year-olds, 8-year-olds, and adults, respectively. Taken together, these indexes indicated that two-dimensional Euclidean models gave a good fit to the dissimilarity values. Because the addition of a third dimension did not increase the fit, the two-dimensional solutions were retained.

Locations of the 20 videotaped pictures in the two-dimensional space for each age group are shown in Fig. 2. In each group, the vertical axis represents differences in valence (with pleasant stimuli at the top and unpleasant stimuli at the bottom), whereas the horizontal axis represents differences in arousal (with low-arousal stimuli on the left and high-arousal stimuli on the right).

To provide a quantitative comparison of the MDS solutions for the different age groups, canonical correlation analyses were carried out among the age groups on the average coordinates obtained for each stimulus in the two-dimensional space. Correlation values were high (.994, .976, and .991) and statistically significant ($p < .001$) for each comparison (between 6-year-olds and 8-year-olds, between 6-year-olds and adults, and between 8-year-olds and adults, respectively), indicating that the two dimensions of valence and arousal were shared by all age groups.

Further correlational analyses were performed to test for associations between the mean coordinates of each portrayal along Dimensions 1 and 2 and their mean of valence, arousal, and intensity

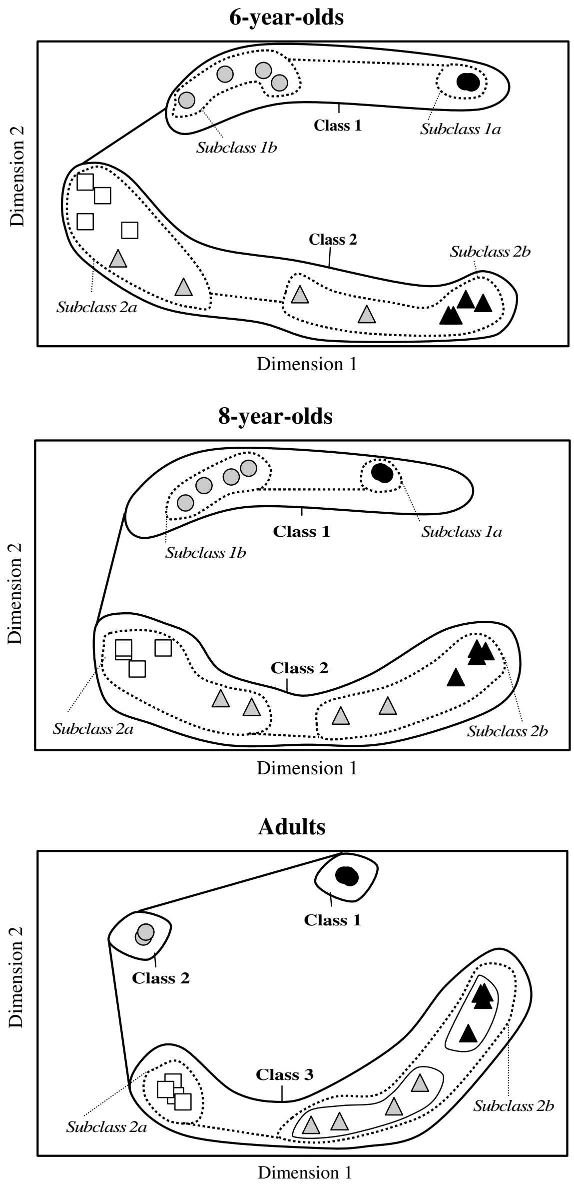

Fig. 2. Two-dimensional MDS solution for happiness (black circle), pleasure (gray circle), anger (black triangle), irritation (gray triangle), and neutrality (white square) portrayals as a function of age group. The main outcomes of the cluster analyses are superimposed onto each MDS solution. Solid black lines represent the main classes, and dashed lines represent subordinate classes. Thin lines in Subclass 2b in the adults' space represent second-level subordinate classes.

ratings collected in the pilot study. These indicated that Dimension 1 was more highly correlated with arousal ratings, $r(18) = .94$, $p < .001$ in 6-year-olds, $r(18) = .96$, $p < .001$ in 8-year-olds, and $r(18) = .87$, $p < .001$ in adults, than with intensity ratings, $r(18) = .61$, $p < .05$ in 6-year-olds, $r(18) = .64$, $p < .05$ in 8-year-olds, and $r(18) = .56$, $p < .05$ in adults. Furthermore, Dimension 1 was not correlated with valence ratings. Dimension 2, on the other hand, was highly correlated with valence ratings, $r(18) = .97$, $p < .001$ in 6-year-olds, $r(18) = .92$, $p < .001$ in 8-year-olds, and $r(18) = .75$, $p < .001$ in adults, but not with arousal or intensity ratings.

Cluster analysis

An additional hierarchical cluster analysis was processed with SPSS 14.0 software (Kinnear & Gray, 1999) to discover any underlying patterns of similarity in the psychological data (Daws, 1996; Kruskal, 1977). This analysis allows a set of objects to be arranged in clusters (classes) to establish a set of object clusters (tree structure) such that the objects within a given cluster are more similar to each other than they are to those in other clusters. Here the cluster analysis would provide a means of exploring the extent to which the hierarchical psychological structure of expressions evolves with age. According to the dimensional approach, the age increase initially would be associated with a clear distinction between positive and negative classes of expressions and later would be associated with a better dissociation between same-valence expressions characterized by different degrees of arousal (e.g., distinguishing anger from irritation within a class of unpleasant expressions). Three hierarchical tree structures were computed based on the dissimilarity matrices obtained for the 6-year-olds, 8-year-olds, and adults. To this end, we applied the average linkage method, using the average distance between all pairs of objects in any given cluster. The correlation between the dissimilarity matrices and their associated proximity matrices (distances between elements in the tree structure) showed that the tree structure represented more than 81% of variance, with $r(398) = .93$, $p < .001$, in each age group. Furthermore, there were close correlations between the proximity matrices of the 6- and 8-year-olds, $r(398) = .95$, $p < .001$, the 8-year-olds and adults, $r(398) = .84$, $p < .001$, and the 6-year-olds and adults, $r(398) = .76$, $p < .001$.

For each age group, the differentiation between classes was achieved by setting a threshold dissimilarity value of 10[2] on the distance scale (ranging from 0 to 25) of the tree structures. This allowed us to identify different partitions within the tree structures: four for the 6-year-olds, four for the 8-year-olds, and six for the adults. In each group, the classes and subclasses were superimposed on the corresponding MDS solution illustrated in Fig. 2. For the 6- and 8-year-olds, two main classes (solid lines) were found; one included pleasant expressions (i.e., happiness and pleasure), whereas the other contained all of the unpleasant (i.e., angry and irritated) and neutral expressions. Four distinct subclasses (dashed lines) discriminated among happiness, pleasure, neutrality, and anger. Irritated expressions were distributed across the neutrality and anger subclasses. In adults, there were three main classes delineating happiness, pleasure, and all other stimuli. There were two subclasses distinguishing neutral from unpleasant expressions as well as two other subordinate, second-level classes (thin lines) differentiating irritation from anger.

Verbalization analysis

A verbalization analysis was carried out to probe the children's semantic abilities. As described above, the production of correct labeling was not expected to improve greatly between 6 and 8 years of age. To test this hypothesis, participants were instructed to provide one label describing the emotion that explained why the portrayals belonged to the same group. The 6-year-olds produced a mean number of 3.44 labels ($SE = 0.19$), the 8-year-olds produced 4.54 labels ($SE = 0.20$), and the adults produced 5.50 labels ($SE = 0.18$). Not surprisingly,[3] there was a significant effect of age, $F(2, 80) = 30.701$, $p < .001$, $\eta^2 = 1$. A significant linear trend, $F(1, 80) = 1641.284$, $p < .001$, indicated that the 6-year-olds'

[2] The threshold value for class separation was chosen so that a single value for all three tree structures (for 6-year-olds, 8-year-olds, and adults) would yield a number of classes that was as close as possible to the number of intended expressions.

[3] The mean number of labels was dependent on the mean number of (sub)classes.

mean number of labels was lower than that of the 8-year-olds, which in turn was lower than that of the adults.

The verbalization analysis consisted in matching the participants' labels to each class and subclass (first level) identified in the cluster analysis. To this end, participants' groupings of portrayals were compared with the classes yielded by the cluster analysis. If the portrayals in an individual participant's grouping included more than 50% of the portrayals in a given class, the label for that grouping was assigned to the class in question. Note that a given label, therefore, could be assigned to more than one class if the participant-defined grouping included at least 50% of the portrayals in more than one class derived from the tree structure. This technique ensured that the labels were assigned in a relatively conservative manner.

The labels that were most frequently assigned to the classes and subclasses as a function of age are summarized in Table 1. This summary was achieved by compiling all of the labels that had been produced at least twice by participants in each of the three age groups. The 6-year-olds produced a total of 29 different labels, the 8-year-olds produced 39 labels, and the adults produced 60 labels. As can be seen in the table, the 6- and 8-year-olds characterized the portrayals in Class 1 as depicting "contentment" and "happiness," whereas Class 2 was predominantly associated with terms designating "sad," "unhappy," "angry," and "irritated" expressions. In Class 1, labels associated with Subclasses 1a and 1b designated different degrees of positive expressions (e.g., "very happy" vs. "little content"), although most of the time the children assigned the same term to both classes (e.g., "happiness"). In Class 2, Subclass 2a was principally associated with sadness, whereas Subclass 2b included labels such as "angry," "unhappy," and "irritated." In adults, Class 1 referred to "happiness," "intense joy," and "enthusiasm," whereas Class 2 was principally associated with "easiness," "contentment," and "pleasure." Class 3 grouped together expressions associated with "expectation" and "neutrality," on the one hand, and "anger" and "irritation," on the other.

A quantitative analysis was performed on the verbalizations by calculating the number of incorrect labels produced by each participant in each age group. The number of incorrect labels was scored by having two "blind" judges select labels that did not match the target expression. The blind judges were

Table 1
Most frequently verbalized groups of similar labels assigned to each class and subclass (first level) of portrayals as a function of age

Class 1		Class 2	
Subclass 1a	Subclass 1b	Subclass 2a	Subclass 2b
6 years			
Content (65.4)	Content (28.6)	Sad (16.3)	Unhappy (40.0)
Very happy (11.5)	Good (14.3)	Angry (14.0)	Irritated (18.3)
Happy (7.7)	Feel good (10.4)	Normal (12.8)	Cross (9.2)
	Tired (9.1)	Not happy (11.6)	Very cross (7.5)
8 years			
Happiness (33.7)	Happiness (16.7)	Sad (15.0)	Angry (41.0)
Content (21.1)	Ease (16.7)	Thinking (13.8)	Irritated (28.9)
Happy (12.6)	Little content (13.9)		
Very happy (8.4)	Happy (11.1)		
	Feel good (11.1)		
	Calm (8.3)		
Adults			
Class 1	Class 2	Class 3	
		Subclass 3a	Subclass 3b
Happiness (48.8)	Easiness (50.8)	Expectation (26.2)	Anger (40.0)
Intense joy (13.0)	Contentment (11.9)	Neutral (10.3)	Irritation (13.3)
Very happy (8.9)	Pleasure (9.5)	No emotion (7.5)	Exasperation (6.7)
Enthusiasm (6.5)	Well-being (6.3)		
	Satisfaction (6.3)		

Note. Mean percentages of labels are indicated in parentheses. Labels were originally produced in French and are presented here in approximate English translation.

instructed to decide whether each label was compatible with the description associated with the intended expression. Note that for neutral stimuli, the blind judges were asked to decide whether the labels were compatible with the description "the person does not feel anything in particular." In each age group, Cohen's kappa coefficient was performed on the number of incorrect labels counted by each judge so as to measure interrater reliability. These coefficients reached .83, .70, and .50 for the 6-year-olds, 8-year-olds, and adults, respectively, indicating high to middle interrater agreement. Given the fairly good interrater reliability, the proportions of incorrect labels were averaged across the blind judges for each expression category and for each age group. A repeated-measures ANOVA, with expression category as a within-participants factor and age group as a between-participants factor, was calculated on the mean proportion of incorrect labels. There was a main effect of age group, $F(2, 80) = 22.225$, $p < .001$, $\eta^2 = .99$, and expression category, $F(4, 320) = 64.418$, $p < .001$, $\eta^2 = 1$. As illustrated in Fig. 3, the interaction between expression categories and age groups was significant, $F(8, 320) = 3.762$, $p < .001$, $\eta^2 = .98$. Post hoc Fisher tests indicated that the mean proportion of incorrect labels was significantly lower ($p < .001$) in adults than in children for irritation (6-year-olds: $M = .51$, $SE = .06$; 8-year-olds: $M = .56$, $SE = .06$; adults: $M = .27$, $SE = .06$), for pleasure (6-year-olds: $M = .35$, $SE = .06$; 8-year-olds: $M = .30$, $SE = .06$; adults: $M = .03$, $SE = .05$), and for neutrality (6-year-olds: $M = .70$, $SE = .07$; 8-year-olds: $M = .71$, $SE = .08$; adults: $M = .31$, $SE = .07$). No significant difference was observed between the 6- and 8-year-olds whatever the expression category.

To control whether verbal skills were associated with emotion processing performance on the free classification task, correlation analyses were performed in each age group between the mean proportion of incorrect labels and the number of groupings. No significant correlation was found in either the children ($r(25) = -.03$ for 6-year-olds, $r(22) = .21$ for 8-year-olds) or the adults, $r(30) = .19$.

Discussion

The objective of this study was to examine children's and adults' abilities to freely categorize and label dynamic bodily/facial expressions of anger, irritation, happiness, pleasure, and neutrality that had previously been controlled for their valence, arousal, intensity, authenticity, and clarity values in a pilot study. Expression categories were differentiated in a two-dimensional space of valence and arousal to control for their respective effects on classification and labeling task performances.

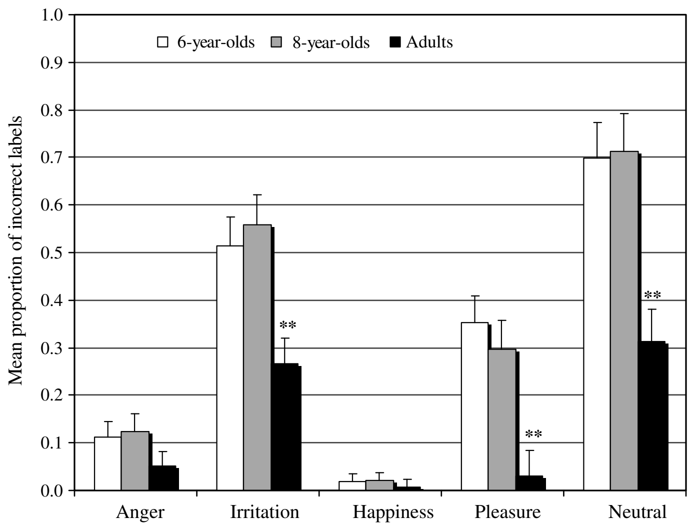

Fig. 3. Mean proportions of incorrect labels as a function of expression categories and age groups.

In the pilot study, the clarity and authenticity ratings were above the median scale value, indicating that all portrayals were judged to be relatively unambiguous and emotionally authentic. Furthermore, adults' performances on the free classification and labeling tasks demonstrated that the intended expressions were accurately discriminated and fairly well labeled. Taken together, these findings provide evidence that the stimuli successfully conveyed the intended expressions.

Interestingly, the results indicated that the arousal and intensity ratings did not probe the same stimulus features; some expressions with a low arousal score (e.g., pleasure) were judged to be as intense as those with a higher arousal score (e.g., happiness). As expected, a low-arousal expression was not necessarily low in intensity and vice versa. The data supported the hypothesis that the arousal ratings probed the degree of energy, whereas the intensity ratings captured the depth of emotion conveyed by the expressions. This sheds some light on what the intensity and arousal dimensions actually reflect and constitutes a good first step in examining how and to what extent some changes in human expressions are associated with changes in energy and emotional intensity.

In the children, the age-related increases in the number of groupings and the number of emotional labels indicated a quantitative improvement but did not exactly mirror the way in which children actually perceived and interpreted intended expressions. Multidimensional and correlation analyses indicated that adults and children judged the similarity between expressions according to their valence and arousal (rather than their intensity). This provides further evidence that these two main dimensions govern emotional responses. Furthermore, it is consistent with Russell's circumplex model of affect and supports the view that younger and older children alike may structure emotions in terms of valence and arousal (Bullock & Russell, 1984, 1985; Russell & Ridgeway, 1983). Our findings also go one step further, demonstrating that children's emotional responses to finely tuned expressions—not merely basic expressions—are organized along these two dimensions. Moreover, although a slightly more refined classification was observed in adults, there were high levels of consistency among the MDS solutions whatever the age group. This means that the structure of the emotional stimuli that emerged from the free classification task was only weakly influenced by age. Importantly, the MDS solutions for the 6- and 8-year-olds were relatively similar, indicating that the ability to perceive salient emotional cues of valence and arousal is already effective as early as 6 years of age. Interestingly, the configuration of expressions in the MDS solutions closely resembled the boomerang-shaped distribution found in the biphasic model developed by Lang and colleagues (1999). This suggests that children, like adults, base their affective judgments on two appetitive and defensive motivational systems.

The cluster analysis indicated that both adults and children placed expressions in distinct classes. The adults created as many (sub)classes as there were intended expressions, discriminating between happiness and pleasure as two distinct classes and assigning irritation, anger, and neutral expressions to a broader class. As for the 6- and 8-year-olds' groupings, each of these age groups produced four (sub)classes. First they distinguished between two broad pleasant–unpleasant classes, and then they divided the former into two subclasses corresponding to intended happiness and pleasure and divided the latter into two subclasses of intended anger and neutral portrayals (irritation was fuzzier). As expected, when children needed to discriminate expressions without drawing on their verbal abilities, they first based their discrimination on a valence dimension in terms of pleasant–unpleasant expressions. This supports the hypothesis that the valence dimension plays a primary role in organizing the emotional domain in children (Bullock & Russell, 1984). Furthermore, the fact that the 6- and 8-year-olds processed pleasure and happiness as two subclasses pertaining to a broader "pleasant" class, whereas adults treated them as two very distinct classes, is consistent with the view that, however expressions come to be categorized, they are first perceived in terms of valence (and arousal) rather than their membership of different categories (Bullock & Russell, 1984). The children's classification was actually slightly more general than that of adults, especially for irritation, which displayed fuzzy boundaries with anger and neutrality. These results suggest that children perceive neutral expressions as rather unpleasant, consistent with previous data indicating that children under 9 years of age tend to perceive either happy or sad emotions when faced with neutral expressions (Durand et al., 2007). Findings also suggest that children did not clearly distinguish between the different degrees of arousal of irritation and anger.

An important question to ask is why expressions of happiness/pleasure were discriminated better than expressions of anger/irritation. The assumption that a positive emotion is recognized better because it does not compete with other positive ones (De Sonneville et al., 2002) cannot account for the current findings. One possible explanation is that happy and pleased expressions are more easily processed because they are most familiar to most children (Batty & Taylor, 2006). Nonetheless, it should be pointed out that adults also displayed better discriminatory powers for positive expressions than for negative ones. This indicates that the boundaries between irritated and angry expressions may be intrinsically less clear than those between pleasure and happiness.

Labeling findings indicated that although the 8-year-olds produced more terms than the younger children, they were no better at labeling expressions accurately. Instead, both the 6- and 8-year-olds tended to assign the same words (''heureux''/''happy'' in 8-year-olds and ''content''/''pleased'' in 6-year-olds) to both positive expressions of happiness and pleasure. The tendency to overgeneralize emotion words to all emotions with similar levels of pleasure and/or arousal has already been observed in 3-, 4-, and 5-year-olds (Bullock & Russell, 1984). The current data extend these previous findings, showing that overgeneralization may continue until approximately 8 years of age, when expressions are very close in terms of valence. Furthermore, the children often used the label ''sad'' to describe a subclass that included neutral expressions, consistent with a previous report (Durand et al., 2007). It could be argued that the neutral expressions in our study were slightly ambiguous stimuli because they were judged to be slightly more negative and less authentic. This pattern may support the view that no facial expression is completely without traces of affect. Even so, we found that the low arousal and rather intermediate valence levels of our neutral stimuli matched the operational definition of neutrality in Lang and colleagues' (1999) motivational model. It is also interesting to point out that the adults successfully discriminated the neutral stimuli, labeling them as ''expectation,'' ''neutral,'' or ''no emotion'' but never as ''sad.''

Further research is needed to clarify children's tendency to assign emotional significance to neutral expressions. The current findings also provided evidence that children can label prototypical expressions of happiness and anger expressions as accurately as adults, although they were less efficient at labeling expressions of pleasure, irritation, and neutrality correctly. This is in line with the view that children may need prototypical elements to effortlessly identify and label emotions (Balconi & Carrera, 2005). Furthermore, this is consistent with the fact that the first affective terms most often used by children are ''happy,'' ''angry,'' and ''sad'' (Widen & Russell, 2003), and it confirms that, in child development, primary expressions are easier to label and conceptualize than are less typical ones.

Finally, it is worth pointing out that the 6-year-olds preferentially used adjectives to label expressions, the adults mostly employed nouns, and the 8-year-olds used a mixture of both. This development in speech between 6 and 8 years of age was not associated with any improvement in assigning correct labels to expressions. This provides interesting insights into how language skills and emotions may be related in development and calls for further investigation.

Current findings also showed that the number of groupings made through free classification was not related to the ability to label expression categories correctly. In other words, even if children did not succeed in assigning labels accurately, they were still able to perceive nuances among the expressions. Second, although the 8-year-olds varied their labels more than their younger counterparts, their abilities to label expressions correctly were no better. Third, children perceived emotional similarities in nearly the same way as adults did but were rather inaccurate in their production of labels for nonprototypical expressions (e.g., pleasure, irritation). Taken together, these findings confirm that children's abilities to perceive (dis)similarities among emotional cues do not depend on their semantic lexical skills. Instead, they confirm that improvements in producing the correct labels take place at a slow pace. This is in line with Bullock and Russell's data indicating that children may organize expressions along the valence and arousal dimensions, using broadly applied archetypical labels (overgeneralization) before establishing label categories that match adult standards.

Findings also indicated that older children did not classify expressions any differently from younger children. The fact that 6- and 8-year-olds distinguish expressions in a similar fashion can be explained by their ability to use valence and arousal cues that have been found to be easily perceived by children as young as 3 years (Bullock & Russell, 1984, 1985).

Limitations of the study and future directions

The testing procedure we used for young participants was designed to minimize the direct potential influence of the experimenter's behavior (e.g., tone of voice, body posture, facial expressions) on the children's emotional responses, thereby providing some guarantee that our data would not be affected by direct behavioral cueing. However, despite these precautions, we cannot exclude the possibility that some subtle and unconscious behavioral cueing on the part of the experimenter may have influenced the children's decision making. In future research, this question will need to be addressed by using a computing device so that children can perform the task autonomously while the experimenter remains oblivious to the emotional content of the stimuli. An interface masking device could be used to allow children to perform the task on a separate touch screen while the experimenter works on another screen using icons representing the stimuli without conveying their actual emotional content. This device could also be used to assess the abilities of children under 6 years of age to categorize and label dynamic expressions that are more or less emotionally similar. This would yield a broader description of the development of these emotional skills.

Conclusion

The current study sought to investigate the abilities of 6-year-olds, 8-year-olds, and adults to categorize and label dynamic bodily/facial expressions that had been controlled for their level of clarity, valence, arousal, intensity, and authenticity. Taking into account the fact that bodily and facial expressions constitute two major nonverbal tools for emotional communication, the current study was designed to make the expressions as life-like as possible. One important issue of the study was to highlight the fact that children and adults structure expressions in systematic and broadly similar ways. Findings confirmed that children's emotional responses were based on the two main dimensions of valence and arousal and that their abilities to perceive dissimilarities among finely tuned expressions did not rely on their semantic lexical abilities. The children's small improvement in correct label production and their relatively nuanced perception of (dis)similarities among expressions support the idea that they may initially interpret emotions dimensionally rather than in terms of clear-cut categories as adults do. This highlights the relevance of both dimensional (i.e., MDS) and categorical (i.e., cluster analysis) methodologies to the investigation of emotional perception and the importance of using dynamic stimuli incorporating face and body movements to reveal the psychological mechanisms behind expression processing in real life. In future investigations, dimensional and categorical methodologies featuring dynamic emotional material could be used to take behavioral research beyond the traditional methods of assessment (e.g., accurate/inaccurate categories, matching/recognition) and stimulus display (e.g., static picture, face separated from body) that may underestimate participants' abilities to process emotions.

Acknowledgments

This research was funded by the ANR "Emotion" program (Project NO05-4_43763). We thank the teachers and principals of the Jules Julien, Rangueil, Ricardi, and Saint Exupère schools as well as all of the children and adults who took part in the study. We are grateful to Clara Martinot and Valérie Tartas for their comments and suggestions, and we also express our gratitude to T. Bänziger, H. Pirker, and K. R. Scherer for granting us permission to use the GEMEP corpus.

Appendix A

In the pilot study, 14 graduates or doctoral-level participants (8 women and 6 men) with normal/corrected vision (24–50 years of age, $M = 36$ years), who were not involved in the experimental session, were asked to perform two sets of ratings on a Macintosh computer running PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993).

First, "emotional clarity" was rated on a 10-point scale (0 = *pas du tout*/not at all, 9 = *tout à fait*/completely) for each portrayal. The instruction was to assess the extent to which the videotapes clearly portrayed the following descriptions (partially taken from the instructions given to the actors portraying the expressions): (a) for happiness: "the person feels transported by a splendid thing, he/she expresses an exalted joy"; (b) for pleasure: "the person feels an exceptional well-being, he/she expresses pleasure"; (c) for anger: "the person feels violent dissatisfaction, he/she expresses great anger"; (d) for irritation: "the person feels a strong discontentment, he/she expresses irritation"; and (e) for neutral stimuli: "the person feels an emotion." This latter instruction was chosen to avoid the possibility of participants judging a stimulus to be "neutral" or "vacant expression" by default each time it failed to match the descriptions of happiness, pleasure, anger, or irritation. The videotapes with both the maximum (minimum in neutral condition) mean (range = 5.2–6.8) and minimum standard deviation values (range = 1.9–2.5) were retained, resulting in the selection of 40 expressions.

One week later, each participant performed a further set of ratings of the selected videotapes on the other 10-point scales. They were instructed to assess the valence (0 = *désagréable*/unpleasant, 9 = *agréable*/pleasant), arousal (0 = *faible énergie*/low energy, 9 = *forte énergie*/high energy), intensity (0 = *peu intense*/not very intense, 9 = *très intense*/very intense), and authenticity (0 = *pas authentique*/inauthentic, 9 = *authentique*/authentic) for each stimulus.

Appendix B

The two sets of stimuli were organized according to expression category ("ANG" for anger; "HAP" for happiness; "IRR" for irritation; "NEU" for neutrality, and "PLE" for pleasure) and to actor ("O" for old, "Y" for young, "F" for female, and "M" for male). The mean clarity, valence, arousal, intensity, and authenticity ratings (and standard deviations in parentheses) are provided for each stimulus.

Stimuli (Set 1)	Clarity (max = 9)	Valence (max = 9)	Arousal (max = 9)	Intensity (max = 9)	Authenticity (max = 9)
ANG01_OF	8.79 (0.43)	0.33 (0.65)	8.86 (0.36)	8.86 (0.36)	8.29 (1.14)
ANG01_OM	7.21 (2.46)	0.69 (0.85)	8.50 (0.85)	8.14 (1.51)	5.43 (3.16)
ANG01_YF	5.43 (1.87)	1.31 (1.03)	6.71 (1.27)	7.36 (1.08)	5.14 (2.48)
ANG01_YM	5.86 (2.91)	1.31 (1.25)	8.00 (0.96)	7.86 (1.29)	5.86 (2.07)
HAP02_OF	7.07 (1.62)	8.15 (0.80)	7.00 (1.47)	7.64 (1.08)	6.71 (2.05)
HAP04_OM	7.50 (1.34)	8.00 (0.91)	7.14 (1.51)	7.36 (1.45)	5.00 (3.01)
HAP01_YF	8.29 (0.91)	7.92 (1.26)	7.29 (1.20)	7.71 (1.43)	6.93 (1.82)
HAP01_YM	7.00 (1.41)	7.15 (1.41)	6.93 (1.38)	7.57 (1.22)	5.14 (2.63)
IRR02_OF	7.57 (1.09)	1.69 (1.18)	4.00 (1.80)	6.07 (1.49)	6.14 (2.07)
IRR05_OM	4.14 (2.54)	2.69 (0.85)	3.00 (1.36)	3.93 (1.27)	4.57 (2.38)
IRR03_YF	3.50 (1.79)	2.38 (1.50)	1.86 (1.41)	3.36 (1.69)	4.29 (2.30)
IRR03_YM	6.36 (2.56)	2.23 (0.93)	5.64 (1.50)	6.07 (1.14)	6.29 (1.82)
NEU02_OF	5.21 (2.39)	2.85 (1.52)	0.23 (0.44)	1.64 (2.13)	4.14 (2.57)
NEU01_OM	5.64 (2.17)	3.08 (0.95)	0.57 (0.76)	1.64 (1.78)	4.64 (2.31)
NEU01_YF	3.25 (2.09)	3.62 (1.71)	0.54 (0.66)	1.43 (1.28)	4.93 (2.06)
NEU05_YM	4.90 (2.37)	4.77 (1.48)	0.79 (0.80)	1.71 (0.91)	4.79 (1.97)
PLE01_OF	5.43 (1.74)	7.08 (1.75)	1.86 (1.10)	5.21 (2.46)	5.43 (2.10)
PLE02_OM	4.86 (1.92)	6.00 (1.29)	1.36 (0.01)	3.43 (1.91)	5.64 (2.24)
PLE02_YF	5.07 (1.64)	5.69 (1.44)	1.29 (1.33)	3.14 (2.14)	3.93 (2.23)
PLE01_YM	5.86 (2.03)	6.54 (0.97)	2.64 (1.08)	4.07 (1.73)	6.00 (2.00)
ANG03_OF	8.21 (1.37)	0.57 (0.97)	8.71 (0.73)	8.77 (0.44)	7.92 (1.12)
ANG02_OM	7.21 (1.67)	0.92 (1.08)	8.50 (0.94)	8.15 (0.99)	5.50 (2.71)
ANG03_YF	7.64 (1.69)	0.77 (0.93)	7.71 (1.20)	7.69 (0.75)	6.71 (2.05)
ANG02_YM	4.57 (2.31)	1.69 (1.18)	5.71 (1.38)	6.36 (1.39)	5.43 (2.38)
HAP03_OF	6.29 (1.68)	7.77 (0.73)	4.93 (1.73)	6.57 (1.22)	5.93 (1.98)

(*continued on next page*)

Appendix B (*continued*)

Stimuli (Set 2)	Clarity (max = 9)	Valence (max = 9)	Arousal (max = 9)	Intensity (max = 9)	Authenticity (max = 9)
HAP15_OM	6.43 (1.91)	7.62 (1.50)	7.07 (1.64)	7.50 (1.65)	4.79 (2.49)
HAP05_YF	8.00 (1.84)	7.69 (1.70)	8.07 (1.07)	8.00 (1.04)	6.93 (2.09)
HAP02_YM	5.79 (1.85)	7.23 (1.24)	6.71 (1.33)	7.29 (1.27)	6.29 (1.54)
IRR03_OF	6.71 (1.49)	2.23 (1.69)	4.36 (1.78)	5.64 (1.74)	5.07 (2.43)
IRR10_OM	4.07 (2.02)	2.54 (1.20)	2.79 (0.97)	3.36 (1.74)	5.79 (1.72)
IRR04_YF	3.14 (2.14)	2.23 (1.42)	2.64 (1.82)	2.79 (1.42)	4.14 (2.57)
IRR08_YM	6.36 (1.86)	2.38 (0.87)	4.57 (1.45)	5.43 (1.55)	5.57 (2.41)
NEU06_OF	5.43 (2.06)	3.31 (1.32)	0.57 (0.65)	1.71 (2.16)	5.21 (2.67)
NEU05_OM	5.29 (2.30)	3.62 (1.39)	0.79 (0.80)	2.64 (1.95)	5.07 (2.30)
NEU03_YF	3.86 (1.74)	3.77 (1.17)	0.86 (1.10)	1.50 (1.02)	4.21 (2.22)
NEU07_YM	4.93 (2.34)	3.46 (0.97)	0.54 (0.78)	1.71 (1.59)	5.07 (2.02)
PLE02_OF	6.36 (1.91)	7.08 (1.44)	1.50 (1.22)	5.29 (2.13)	5.71 (2.27)
PLE04_OM	4.93 (2.02)	6.31 (1.75)	1.71 (0.83)	3.29 (1.68)	4.21 (1.89)
PLE04_YF	6.21 (1.53)	6.77 (1.69)	2.07 (1.27)	4.43 (1.55)	4.71 (2.05)
PLE02_YM	5.14 (2.38)	6.92 (2.06)	2.43 (1.34)	4.64 (2.24)	4.43 (2.98)

References

Abelson, R. P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology, 65*, 546–554.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669–672.

Balconi, M., & Carrera, A. (2005). Il lessico emotivo nel decoding delle espressioni facciali [Lexicalization and conceptualization of emotional expressions in school-aged children: The effect of contextual cues]. *Psychofenia, 12*, 17–35.

Bänziger, T., Pirker, H., & Scherer, K.R. (2006, May). *GEMEP–Geneva Multimodal Emotion Portrayals: A corpus for the study of multimodal emotional expressions*. Paper presented at the Fifth International Conference on Language Resources and Evaluation, Genoa, Italy.

Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental Science, 9*, 207–220.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and excerpts' duration. *Cognition and Emotion, 19*, 1113–1139.

Boone, R. T., & Cunningham, J. G. (1998). Children's decoding of emotion in expressive body movement: The development of cue attunement. *Developmental Psychology, 34*, 1007–1016.

Boyatzis, C. J., Chazan, E., & Ting, C. Z. (1993). Preschool children's decoding of facial emotions. *Journal of Genetic Psychology, 154*, 375–382.

Bullock, M., & Russell, J. A. (1984). Preschool children's interpretation of facial expressions of emotion. *International Journal of Behavioral Development, 7*, 193–214.

Bullock, M., & Russell, J. A. (1985). Further evidence on preschoolers' interpretation of facial expressions. *International Journal of Behavioral Development, 8*, 15–38.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology, 13*, 699–745.

Camras, L., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9*, 84–94.

Caron, R. F., Caron, A. J., & Myers, R. S. (1982). Abstraction of invariant face expressions in infancy. *Child Development, 53*, 1008–1015.

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

Custrini, R. J., & Feldman, R. S. (1989). Children's social competence and nonverbal encoding and decoding of emotions. *Journal of Clinical Psychology, 18*, 336–342.

Daws, J. T. (1996). The analysis of free-sorting data: Beyond pairwise co-occurrences. *Journal of Classification, 13*(1), 57–80.

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology, 24*, 200–213.

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J-Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology, 97*, 14–27.

Ekman, P. (1982). *Emotion in the human face*. Cambridge, UK: Cambridge University Press.

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

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expression by neonates. *Science, 218*, 179–181.

Gosselin, P. (1995). Development of facial expressions recognition in children. *Canadian Journal of Behavioral Science, 27*(3), 107–119.

Gosselin, P. (2005). The Decoding of the facial expressions during childhood. *Canadian Psychology, 46*(3), 126–138.

Gosselin, P., & Larocque, C. (2000). Facial morphology and children's categorization of facial expressions of emotions: A comparison between Asian and Caucasian faces. *Journal of Genetic Psychology, 161*, 346–358.

Gray, J. M., Young, A. W., Barker, W. A., Curtis, A., & Gibson, D. (1997). Impaired recognition of disgust in Huntington's disease gene carriers. *Brain, 120*, 2029–2038.

Green, R. S., & Cliff, N. (1975). Multidimensional comparisons of structures of vocally and facially expressed emotion. *Perception and Psychophysics, 17*, 429–438.

Herba, C. M., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry, 47*, 1098–1106.

Izard, C. E. (1997). Emotions and facial expressions: A perspective from differential emotions theory. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 57–77). Cambridge, UK: Cambridge University Press.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*, 759–763.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage, 18*, 156–168.

Kinnear, P. R., & Gray, C. D. (1999). *SPSS for Windows made simple*. Hove, UK: Psychology Press.

Kruskal, J. (1977). The relationship between multidimensional scaling and clustering. In J. Van Ryzin (Ed.), *Classification and clustering* (pp. 17–44). New York: Academic Press.

Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (1999). International Affective Picture System (IAPS): Instruction manual and affective ratings. Technical Report No. A-4, Center for Research in Psychophysiology, University of Florida.

Lemay, G., Kirouack, G., & Lacouture, Y. (1995). Spontaneous static and dynamic facial expressions: A comparison of categorical and dimensional judgments. *Canadian Journal of Behavioral Science, 27*(2), 125–139.

Montague, D. P. F., & Walker-Andrews, A. S. (2001). Peekaboo: A new look at infants' perception of emotion expressions. *Developmental Psychology, 27*, 826–838.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383*, 812–815.

Osgood, C. E. (1966). Dimensionality of the semantic space for communication in facial expressions. *Scandinavian Journal of Psychology, 7*, 1–30.

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., et al (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society of London B, 265*, 1809–1817.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102–141.

Russell, J. A., & Ridgeway, D. (1983). Dimensions underlying children's emotion concepts. *Developmental Psychology, 19*, 795–804.

Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4- to 6-month-old infants. *Developmental Psychobiology, 25*, 411–425.

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., et al (1996). Loss of disgust: Perception of faces and emotions in Huntington's disease. *Brain, 119*, 1647–1665.

Van Meel, J., Verburgh, H., & DeMeijer, M. (1993). Children's interpretation of dance expressions. *Empirical Studies of the Arts, 11*, 117–133.

Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: The intersection of perceptual and semantic categories. *Acta Paediatrica, 89*, 836–845.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (in press). Happy, sad, scary, and peaceful musical excerpts for research on emotions. Cognition and Emotion.

Walker-Andrews, A. S. (1998). Emotions and social development: Infants' recognition of emotions in others. *Pediatrics, 102*, 1268–1271.

Werhle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology, 78*, 105–119.

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology, 39*, 114–128.