

Facial appearance and impressions of 'credibility': The effects of facial babyishness and age on person perception

Jaume Masip, Eugenio Garrido, and Carmen Herrero

University of Salamanca, Spain

The babyface overgeneralization effect is perceiving that people whose facial features resemble those of children have childlike traits, and treating them accordingly. This experiment sought to replicate the US findings with a South-European sample, to examine the impact of facial maturity on impressions of truthfulness, and to examine the influence of age on person perception. Three-hundred and twenty-four Spanish undergraduates were shown a photograph and had to rate it on a series of behavioural-tendency and trait scales measuring honesty, truthfulness, strength, dominance, intelligence, naivety, and warmth. The photographs were babyfaced, intermediate, and mature faced computer-manipulated versions of three pictures of the same individual at three different ages. Results indicate that the experimental manipulations significantly affected most of the dependent variables. Babyfaced individuals were perceived as the most truthful, and children as the most deceitful. However, when the deceit concerned a sexual abuse allegation, children were rated as the least deceitful. These results support the existence of (a) the babyface overgeneralization effect, (b) the stereotype that children are unreliable witnesses, and (c) the belief that children never lie about sexual abuse offenses. They also suggest that facial babyishness and age may be static perceived deception cues that may account for the demeanour bias found in nonverbal research on the detection of deceit.

L'effet de sur-généralisation du visage d'apparence enfantine (babyface) consiste à percevoir les gens dont les caractéristiques faciales ressemblent à celles des enfants comme possédant des traits enfantins ainsi qu'à les traiter en accord avec cette perception. Cette expérience cherchait à répliquer les résultats obtenus aux États-Unis avec un échantillon de l'Europe du Sud, d'examiner l'impact de la maturité faciale sur l'exactitude des impressions et d'examiner l'influence de l'âge sur la perception de la personne. Trois cent vingt-quatre universitaires espagnols ont reçu une photographie qu'ils avaient à évaluer sur une série d'échelles de tendances comportementales et de traits mesurant l'honnêteté, la sincérité, la force, la dominance, l'intelligence, la naïveté et la chaleur. Les photographies représentaient une figure d'apparence enfantine, une figure intermédiaire et une figure mature, lesquelles étaient des versions manipulées informatiquement de trois photographies d'un même individu à trois âges différents. Les résultats indiquent que les manipulations expérimentales affectent significativement la plupart des variables dépendantes. Les individus ayant un visage d'apparence enfantine furent perçus comme plus sincères et les enfants comme plus menteurs. Cependant, quand le mensonge concernait une allegation d'abus sexuel, les enfants étaient évalués comme les moins menteurs. Ces résultats soutiennent l'existence de: (a) l'effet de sur-généralisation du visage d'apparence enfantine, (b) le stéréotype que les enfants sont des témoins peu fiables et (c) la croyance que les enfants ne mentent jamais à propos de l'abus sexuel subi. Ils suggèrent aussi que le visage enfantin et l'âge peuvent être des indices statiques de mensonges perçus qui peuvent contribuer au biais d'attitude trouvé dans la recherche non verbale sur la détection du mensonge.

Correspondence concerning this article should be sent to Jaume Masip, Department of Social Psychology and Anthropology, University of Salamanca, Facultad de Psicología, Avda. de la Merced, 109-131, 37005 Salamanca, Spain (E-mail: jmasip@usal.es).

The authors wish to express their gratitude to two anonymous reviewers for their helpful comments on an earlier draft of this manuscript, and to Josep Llevat, Pere Palleja, and Amparo Matilla, who provided us with some series of photographs that made it possible to conduct this research. Furthermore, Amparo Matilla also participated as research assistant during the early phases of this experiment. The computer artist Juan A. Muñoz-Hernández was in charge of manipulating the photographs. Technical and IT assistance were kindly provided by our friends and colleagues Jesús Herrero and Jose M. Arana, and we are indebted to them.

The research reported here was supported by the Junta de Castilla y León, Programa de Apoyo a Proyectos de Investigación, Ref. SA52/00B.

El efecto de sobregeneralización del envejecimiento facial consiste en percibir a las personas cuya apariencia facial es similar a la de los niños como poseedores de rasgos infantiles y en tratarlas de forma acorde. En este experimento se pretendió replicar los hallazgos estadounidenses con una muestra sureuropea, examinar el impacto de la madurez facial sobre las impresiones de veracidad y honestidad, e investigar la influencia de la edad sobre la percepción de personas. Trescientos veinticuatro estudiantes españoles recibieron una fotografía y tuvieron que evaluarla en una serie de escalas de tendencias conductuales y de características de personalidad que medían honestidad, veracidad, fortaleza, dominancia, inteligencia, ingenuidad y calidez. Las fotografías eran de la misma persona en tres momentos de su vida, pero habían sido modificadas por ordenador para hacerlas facialmente envejecidas, intermedias o maduras. Las personas envejecidas se percibieron como las más veraces, y los niños como los más mentirosos. Sin embargo, cuando la mentira consistía en una alegación de abuso sexual, los niños fueron evaluados como los *menos* mentirosos. Estos resultados apoyan la existencia de (a) el efecto de sobregeneralización del envejecimiento facial, (b) el estereotipo de que los niños son testigos poco fiables, y (c) la creencia de que los niños nunca mienten sobre delitos de abuso sexual. Asimismo, sugieren también que el envejecimiento facial y la edad pueden ser claves percibidas del engaño de naturaleza estática que pudieran dar cuenta del llamado *demeanor bias* que se ha hallado en la investigación no-verbal sobre la detección del engaño.

INTRODUCTION

Research into the nonverbal detection of deceit has focused on a number of topics. Chief among them are the explorations into how well people can lie or detect deception (Masip, Garrido, & Herrero, 2002), as well as those into the actual cues to deceit (e.g., DePaulo, Lindsay, Malone, Muhlenbruck, Charlton, & Cooper, 2003). However, deception researchers have also explored people's beliefs about the cues to deceit and the perceived deception cues. The latter are those indicators that people actually use to make veracity judgments (see Zuckerman, Koestner, & Driver, 1981).

Implicitly, the conceptualization of the perceived deception cues comprises two types of information conveyed by the sender: static cues and dynamic cues. However, research conducted so far has focused almost exclusively on the dynamic cues (e.g., body movements, vocal characteristics, etc.), neglecting any exploration into static cues such as dress, physical appearance, etc. In view of extant research showing that these characteristics have a strong influence on a series of social judgments (e.g., Damhorst, 1990; Langlois, Kalakanis, Rubenstein, Larson, Hallam, & Smoot, 2000; Zebrowitz, 1997), it is imperative to examine whether they also influence judgments of deceptiveness. In this experiment the influence of a static cue such as facial appearance on impressions of photographed people's sincerity was examined. In addition, work under way is examining whether facial appearance also influences observers' credibility judgments in a lie-detection task.

In 1979, Zuckerman, DeFrank, Hall, Larrace, and Rosenthal found what they termed a

demeanor bias: some senders were consistently judged as honest and some as deceptive, regardless of whether they lied or told the truth. Later research has confirmed the existence of a *demeanor bias* (e.g., Bond, Kahler, & Paolicelli, 1985). The traditional explanation for this bias maintains that it is caused by certain personality traits or social skills of the sender that influence his or her perceptible *demeanor*, which, in turn, has an effect upon the observers' ratings. However, an alternative (though not necessarily rival) explanation is that the senders' appearance may influence the observers' credibility judgments. In other words, the senders' appearance may contain perceived static indicators of deceit, so accounting for the *demeanor bias*.

An aspect of physical appearance that might have an influence on observers' judgments of credibility is the senders' facial appearance, since it has already been shown to influence other kinds of social judgments. A recent exploratory study examining the impact of facial appearance upon the receivers' veracity judgments was conducted by Masip, Garrido, and Herrero (2003, Experiment 2). However, although the several faces used by those researchers differed significantly in age and attractiveness, they were unable to find support for their claims that the senders' facial appearance would affect the observers' judgments or their accuracy in identifying truths or lies. Masip et al. suggested some possible explanations for their results, and stressed the need to manipulate certain characteristics of the facial stimuli that, according to the existing research, influence a number of social judgments, to examine whether this influence extends to the perceived credibility of the sender or target-person. This has been done in the present experiment.

Two facial features that have been shown to influence the perceivers' social judgments are physical attractiveness (e.g., Berscheid & Walster, 1974; Dion, Berscheid, & Walster, 1972; Feingold, 1992; Langlois et al., 2000; Zebrowitz, 1997) and facial babyishness (e.g., Berry & McArthur, 1985, 1986; Montepare & Zebrowitz 1998; Zebrowitz, 1997). As regards honesty, a number of studies have shown that attractive faces are perceived as more honest than unattractive faces (e.g., Berry, 1990, 1991; Berry & Brownlow, 1989; Berry & McArthur, 1985; McArthur & Apatow, 1983-84; Zebrowitz & Montepare, 1992; Zebrowitz, Montepare, & Lee, 1993; Zebrowitz, Voinescu, & Collins, 1996). This finding is consistent with the attractiveness-halo *effect* (e.g., Berscheid & Walster, 1974) or the what-is-beautiful-is-good stereotype (Dion et al., 1972): Those individuals who are physically attractive are believed to possess additional positive personal characteristics such as honesty (see Langlois et al., 2000, for a recent review and meta-analyses). In addition, targets' facial babyishness has also been shown to increase social perceivers' honesty ratings. This latter finding can be explained in terms of the *babyface* overgeneralization *effect*.

The babyface overgeneralization effect

As the human being grows, it experiences certain anatomical changes that are caused either by the maturation process itself or by the force of gravity. Some of these changes affect the cranium and face, making the prototypical facial appearance of an adult quite different from that of a baby (e.g., Enlow, 1982; Montepare & Zebrowitz, 1998; Zebrowitz, 1997). Some of these changes are summarized below:

Cranial structure. During the growth process there is a relative rising of the location of features on the vertical plane of the face (Todd & Mark, 1981); that is, the eyes-nose-mouth set is placed on a higher facial location in adults than in children. Therefore, a prototypical infant's face has a larger forehead and a smaller chin than a prototypical adult's face. When seen in profile, an infant has a relatively receding chin and a relatively prominent forehead (Todd, Mark, Shaw, & Pittenger, 1980); these characteristics are progressively attenuated during the growth process as a result of, among other factors, the jaw enlargement produced by the development of the teeth.

Eyes. The eyes in an infantile face are relatively larger than those in an adult face. The reason is that the eyes grow very little during the maturation process, so in an adult face they are relatively smaller than in an infant's face (Montepare & Zebrowitz, 1998; Zebrowitz, 1997).

Eyebrows. Infants' eyebrows are thinner and are located at a higher position than adults' (Guthrie, 1976, cited by Berry & McArthur, 1986).

Nose. Babies have a pug nose, relatively wider and shorter than an adult's nose (Enlow, 1982). It is not until puberty that the nasal bridge enlarges, especially among males (e.g., Zebrowitz, 1997).

Ears. Infants also have shorter ears than adults (e.g., McArthur & Apatow, 1983-84).

Lips. Infants' lips are less thin than adults' (Zebrowitz, 1997). Babies have a smaller mouth, with lips that are relatively larger than those of adults.

Cheeks. Babies have round, chubby cheeks. On the contrary, owing to their jaw enlargement, normally mature faces are not round but angular or strong.

Other characteristics. In addition to the above, babies usually have skin that is softer and lighter in colour than that of adults. Their hair and eyes also tend to be light-coloured, and their pupils relatively larger than those of adults.

Accurately identifying a person's age has an adaptive value. For example, babies are more vulnerable, naive, dependent, etc. than older individuals. In order for babies to survive, these internal characteristics must be quickly and accurately perceived by more competent, older individuals capable of looking after the babies and meeting their needs. There is evidence that these characteristics of babies are revealed in the very facial features listed above (Zebrowitz, Fellous, Mignault, & Andreoletti, 2003). Therefore, given their adaptive value, people perceive facial features easily (see Montepare & Zebrowitz's, 1998, 2002 reviews), attribute the correct meaning to them (e.g., vulnerability and naivety), and behave accordingly (look after the baby, nurture and educate him or her, etc.) (see, e.g., Montepare & Zebrowitz, 1998; Zebrowitz, 1997). This favours the baby's survival and, as a result, that of the human species. However, since this process is triggered by the morphological features of babies rather than by the babies

themselves, if these features are displayed by something that is not a baby (e.g., an adult) they will generate the same process. This phenomenon is called the babyface overgeneralization effect (e.g., Zebrowitz, 1997; Zebrowitz et al., 2003). The tendency to respond to babies' facial qualities is so strong that responses are overgeneralized to aged faces that nevertheless contain certain babyish features.

Consistent with the babyface overgeneralization effect, research has shown that adult individuals, whose facial characteristics resemble those of babies, are attributed traits similar to babies. Thus, for example, babyfaced adults are perceived as weaker than their mature-faced counterparts in the social and intellectual domains. They are also perceived as warmer, more approachable, more easygoing, kinder, etc. than mature-faced adults (see Montepare & Zebrowitz's, 1998, review, pp. 112–116). This has important social consequences in areas such as interpersonal relationships, job and promotion opportunities, and attribution of guilt and sentence severity in court (Montepare & Zebrowitz, 1998, pp. 124–134).

One characteristic that has been targeted by face researchers is the "honesty" that perceivers attribute to more or less babyfaced stimulus-persons. In general, babyfaced people depicted in pictures or photographs are perceived as more honest than those relatively mature faced. For instance, Zebrowitz et al. (1996) found that babyfacedness showed a strong, positive correlation with perceived honesty in puberty, adulthood, the 30s and the 50s, and predicted perceived honesty in childhood, puberty, adolescence, and the 30s, with all other variables controlled. Berry and McArthur (1985), Berry and Brownlow (1989), and Berry (1990) reported significant positive correlations between babyfacedness and impressions of the honesty of the stimulus-individuals they used in their studies. It was also Berry (1991) who found that facial babyishness increased ratings of sincerity, a composite of perceived honesty, straightforwardness, warmth, naivety, and kindness. Zebrowitz and Montepare (1992) discovered that the influence of babyfacedness upon honesty judgments extended across the life span, especially among males. Zebrowitz et al. (1993) found significant correlations between the facial babyishness of white, black, and Korean targets and honesty impressions of white, black, and Korean observers. In addition, research conducted by Brownlow (1992; Brownlow & Zebrowitz, 1990) indicates that babyfaced individuals are perceived as more trustworthy than mature-faced individuals. All these results are

consistent with the above *babyface overgeneralization* effect.

The present experiment

This study is based on McArthur and Apatow's (1983–84) experiment. They used a police Identikit to build two basic male faces and two basic female faces. Each of these four control faces was then modified by manipulating (a) the size of the eyes (increased in the babyish versions, decreased in the mature ones), (b) the length of the nose and ears (shortened in the babyish versions, enlarged in the mature ones), (c) the vertical placement of facial features (lowered in the babyish versions, moved upward in the mature ones), and (d) all these characteristics simultaneously (full babyish or mature faces). The nine versions of each face (control, four babyfaced versions, and four mature-faced versions) were presented to participants, who had to rate them on a series of bipolar trait scales (measuring strength–weakness, dominance–submissiveness, intelligence–unintelligence, naivety–shrewdness[only for female faces], warmth–coldness, and honesty–dishonesty) and a series of behavioural-tendency scales corresponding to the same traits (e.g., "does s/he look like someone who could explain a very difficult theory to you?" for intelligence–unintelligence, etc.). In general, McArthur and Apatow's results supported their hypotheses: With few exceptions, the babyish faces were perceived as weaker, warmer, more submissive, honest, naive, and unintelligent than the mature faces. These results were not due to facial attractiveness or the stimulus-face perceived age.

The present experiment is a replication and extension of McArthur and Apatow's study with a South-European sample. Replicating psychological studies in a given social or cultural environment is a necessary step before making undue cross-cultural generalizations. Although studies conducted later than the one by McArthur and Apatow have shown that the babyface overgeneralization effect is also apparent in Koreans and in African Americans (McArthur & Berry, 1987; Zebrowitz et al., 1993), to our knowledge only one study (Atzwanger & Grammer, 1993) has been conducted using European participants. The hypotheses of the present experiment were in line with the findings of foreign research: It was predicted that babyish faces would be perceived as weaker, warmer, more submissive, and more naive than mature faces. As concerns intelligence, the findings of McArthur and Apatow (1983–84) were

inconsistent across the different facial manipulations. Further analyses revealed that intelligence perceptions were actually under the influence of attractiveness, instead of being under the influence of facial babyishness. Thus, in the present experiment no significant relationship was expected between facial babyishness and intelligence ratings.

In addition to the sociocultural context, there are a number of additional differences between the present study and that of McArthur and Apatow. First and most importantly, extant research has measured the influence of facial babyishness upon honesty impressions. However, the meaning of the English word honesty—very close to sincerity or truthfulness—is not exactly the same as that of its Spanish counterpart, *honestidad*, which refers to a general quality of character indicating moral integrity, virtue or uprightness¹. Therefore, to properly examine whether facial babyishness has any effect on Spanish perceivers' impressions of the target-participants' truthfulness, truthfulness–deceitfulness scales had to be used rather than honesty–dishonesty scales. Hence, a trait scale and a behavioural-tendency scale to measure perceived truthfulness–deceitfulness were included, while the original McArthur and Apatow's (1983–84) honesty–dishonesty scales were retained as well. The predictions were that babyish faces would be judged as more truthful and honest than less babyish faces.

Second, while McArthur and Apatow (1983–84) controlled the effect of age of the stimulus-person, in this experiment age was manipulated and entered as an independent variable in the analyses. The importance that has recently been given to child witnesses' testimony, especially in sexual abuse cases (e.g., Bull, 2001; Dent & Flin, 1992; Goodman & Bottoms, 1993; Sachsenmaier & Watson, 1998), justifies the inclusion of this variable in the design. Do perceivers think that children, merely because they are children, do not lie? What do they think about children's truthfulness if the issue at hand is a sexual abuse allegation? Using children's faces made it possible to answer the first question, and asking perceivers the extent to which the person in the photograph would accuse an innocent of having committed a sexual abuse offence made it possible to answer

the second question. The prediction was that children would be considered less inclined to lie in a sexual abuse case than adults. This hypothesis was based on social concern with regard to child sexual abuse. Regarding lying in general, no specific hypothesis was posed: The general tendency to consider that children are more innocent, dependent, and honest than adults suggests that they are probably also considered more truthful than adults. However, this view of children coexists with the notion that they are very imaginative and therefore invent, concoct, and lie. In fact there is evidence that, historically, children have been considered unreliable witnesses (e.g., Baartman, 1992; Ceci & Bruck, 1993; Sachsenmaier & Watson, 1998), and this view still remains among some legal professionals (Alonso-Quecuty, 1996; Brigham & Spier, 1992). Older faces were also used in the present experiment for a number of reasons: (1) the dichotomous distinction between children and adults does not capture the diversity of human growing, maturation, and ageing; (2) in view of the biased and stereotypical attitudes that many people hold towards ageing adults ("ageism"; Butler, 1969; see Nelson, 2002, for a recent compilation), it was important to include in the design not only adult faces but also faces of an old person; (3) in their meta-analysis, Kite and Johnson (1988) found that the negative evaluations of old people were most apparent as regards competence ratings (the differences were smaller when desirable contact and personality traits were measured); and (4) it is interesting to examine the generalizability of the effects of facial babyishness across the life span.

Third, instead of the schematic pictures of a police Identikit originally used by McArthur and Apatow, more realistic computer-manipulated photographs were used in this study. This manipulation also made it possible to obtain babyfaced and mature-faced versions of the same face while keeping other facial characteristics unchanged.

Fourth, while McArthur and Apatow (1983–84) used a within-subjects design, a between-subjects design was used in the present experiment. Indeed, McArthur and Apatow argued that it is unlikely that their respondents guessed the experimental hypotheses, but they also acknowledged

¹According to *The Concise Oxford Dictionary*, *honesty* is "1. being honest. 2. truthfulness," and *honest* means "1. fair and just in character or behaviour, not cheating or stealing," but also "...2. free of deceit and untruthfulness, sincere." Thus, the English word *honesty* is very close to the words *truthfulness*, *sincerity*, or *veracity*. This is not so for the Spanish word for *honesty*, namely *honestidad*. The *Dictionary of the Real Academia Española* defines it as "quality of being honest," and defines *honest* (*honesto*) as "1. Decent or decorous. 2. Modest or shameful. 3. Reasonable, fair. 4. Upright." Thus, the Spanish word for *honesty* is farther from veracity or truthfulness than its English counterpart, being closer to a general quality of character indicating uprightness or goodness.

that using a between-subjects design would have been more cautious in this regard. This would certainly increase the sample size, but to deal with this problem the number of experimental stimuli can be limited. Since one of the aims of the present experiment was to confirm in a different context findings that have been found repeatedly in the US, reducing the number of faces was not necessarily a problem—provided the findings were in line with those of foreign research. Therefore, only one stimulus-person was used. Three photographs of this individual (one depicting him as a child, another one as an adult, and the third one as a somewhat older person) were manipulated to create babyfaced and mature-faced versions.

Fifth, only male faces were used in the present experiment. Since, in general, adult females retain more babyish facial features than males (Bruce & Young, 1998; Friedman & Zebrowitz, 1992), an extremely mature-faced female could perhaps be perceived as grotesque or unnatural. Therefore, the only stimulus-person to be used was a male.

Finally, instead of manipulating individual facial characteristics, a number of facial features were changed simultaneously, thus creating full babyish or full mature faces. Unlike McArthur and Apatow, the present authors were not interested in identifying the individual facial features indicative of babyishness that influence perceivers' impressions, but rather in examining whether the overgeneralization effect is also apparent in a European country, as well in exploring whether it extends to perceptions of sincerity. Therefore, the photograph manipulations always involved various facial features simultaneously.

METHOD

Participants

Three-hundred and twenty-four undergraduate students of psychology at a Spanish university participated in this experiment. Their mean age was 23.22 years (range: 20–38, $SD=3.09$). Most of them were females ($N=273$; there were 45 males, and 6 respondents did not report their gender).

²Later on, some undergraduates had to assess on a scale the age of the persons in these photographs. Nine respondents judged the child's face, 10 the adults' face, and 10 the oldest one. All those who judged the child's face indicated that the person in the photograph was younger than 15. The adult face was judged to be around 30 years old. The oldest face, however, was judged to be only around 50 years old although it had been expected to be judged as somewhat older, around 60 years old.

Procedure

Stimulus photographs

A series of three photographs of a male at different ages (as a child, as an adult, and in his 60s) was searched for. In none of these photos was this person to wear glasses or have facial hair, he could not be bald, and in all three pictures his hairstyle had to be similar, his facial expression had to be more or less neutral, and he had to face the camera. A few series of photographs were obtained, and the series that most closely met these requirements was selected.

A computer artist was contacted and given precise instructions concerning the manipulations he had to make on the original photographs. First of all a neutral, white background was placed in all of them, the stimulus-person was "dressed" with the same clothes in all the photographs, and the soldier's cap the subject wore in the original adult picture was removed². Next the facial maturity modifications were made. A babyfaced and a mature-faced version had to be obtained from each of the three faces (the child's, the adult's and that of the oldest individual). Modifications consisted of proportional increases or decreases of certain facial features that differentiate babies from older individuals: forehead size (the distance from the eyes to the hair), chin length, eyebrow position (distance from the eyes to the eyebrows), eye size, nose (pug vs. straight or aquiline), ear length, lips (full vs. thin) and facial roundness (chubby cheeks vs. large, angular and strong jaw). The direction of the manipulations are specified in Table 1, and some of the resulting photographs are shown in Figure 1.

All faces were printed in black and white on DIN-A4 sheets of paper that permitted a good image quality. The pictures were located in the centre of the paper, and their size was 10 x 12.33 cm.

Experimental session

The data were collected during a practical social psychology class. Participants entered the classroom, where they found a questionnaire and one of the nine photographs (3 age levels x 3 babyishness levels) on each desk. There were 36 respondents allocated to each cell resulting from the crossing of the two independent variables (i.e., facial babyishness and age); that is

TABLE 1
Direction of the manipulations made on the intermediate faces to obtain babyish and mature versions

<i>Facial characteristic manipulated</i>	<i>Babyish</i>	<i>Mature</i>
Forehead	Large	Small
Eyebrows	Raised	Lowered
Eyes	Large	Small
Nose	Pug	"Bridged"
Ears	Short	Long
Lips	Full	Thin
Chin	Short	Large
Facial shape	Round (chubby cheeks)	Angular (large and strong jaw)

to say, each face was judged by 36 people³. Once seated, participants were told that they were going to participate in a person perception study and were instructed to look at the photograph and rate it on the scales in the questionnaire. They

were also asked not to leave any question unanswered⁴.

The questionnaires consisted of a series of 7-point scales to rate the **photographs** on a series of traits and behavioural tendencies. Trait scales were on one side of the sheet of paper, and the behavioural-tendency scales on the other; the location of trait or behavioural-tendency scales on the front or back of the paper was counter-balanced, and a note written at the foot of the front side asked respondents not to turn the page until all the scales on that side had been completed. The traits participants were asked to rate were strength–weakness, dominance–submissiveness, intelligence–unintelligence, naivety–shrewdness, warmth–coldness, honesty–dishonesty, and truthfulness–deceitfulness. The questions were Spanish translations and adaptations of those used originally by McArthur and Apatow, although additional items were added to measure perceived

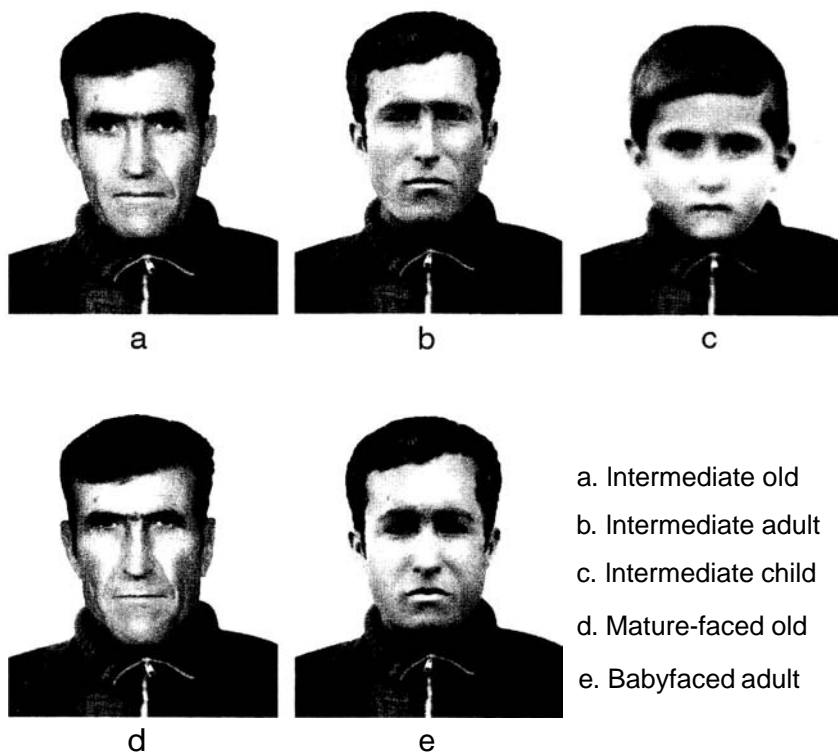


Figure 1. Some of the stimulus photographs.

³Chi-square and loglinear analyses indicated that although there were certainly many more females than males in the sample, this held true in all the experimental conditions: Babyishness \times gender and age \times gender chi-square analyses did not yield any significant effect, respectively: $\chi^2(2) = 0.14$, $p = .934$, $\chi^2(2) = 0.61$, $p = .739$. A stepwise backward hierarchical loglinear analysis where gender and the two independent variables were entered, yielded a model that included only the main effect of gender, partial $\chi^2(1) = 181.55$, $p < .001$, $|\lambda| = 0.87$, $|z| = 11.13$, likelihood ratio $\chi^2(16) = 1.94$, $p = 1.00$; the greatest standardized residual had an absolute value of 0.45.

⁴Despite this instruction, several participants left one or more questions unanswered. This resulted in some missing data in the MANCOVAs, which were calculated later. The number of missing data was not the same in the trait scales as in the behavioural-tendency scales. That is why degrees of freedom for the analysis reported below on trait ratings and behavioural-tendency ratings are not the same.

truthfulness or deceitfulness (a further trait scale and a further behavioural-tendency scale) as well as the extent to which the stimulus-person was perceived as being capable of accusing an innocent of having sexually abused someone else (a further behavioural-tendency scale). All the questions are listed in Table 2.

On the behavioural-tendency scales, 1 was *certainly not*, 4 was *perhaps*, and 7 *certainly*. On the trait scales the same question was repeated for all the scales: "To what extent does this person look...". Endpoints were labelled as indicated in Table 2, with value 4 *intermediate*. In addition, a facial attractiveness measure was included among the trait scales. The question was "How does this person look?" and data were collected on a 7-point scale labelled as in Table 2.

Once respondents had completed the questionnaire, they were thanked for their participation and they left the classroom, leaving the questionnaires and the photographs on the desks so that the experimenter could collect them. Later, during another lecture, they were debriefed.

RESULTS

Overview of data analyses

Two multiple analyses of covariance (MANCOVAs) were conducted through the General Linear Model module of SPSS, one on the trait scale ratings and the other on the behavioural-tendency scale ratings. The independent variables were age (child/adult/old) and facial babyishness (baby-faced/intermediate/mature-faced). Since previous face research has shown that physical attractiveness has a strong influence on social judgments, its effect had to be controlled. Therefore, the attractiveness measure was entered as a covariate in the analyses. Multivariate and univariate effects were examined, and polynomial and simple contrasts were run to discern whether each significant univariate effect had a linear or quadratic trend and to examine the differences between the faces of different age and facial maturity. Mean ratings and most results of the univariate analyses are shown in Tables 3 and 4.

TABLE 2
Scales in the questionnaire

Belzuviourul-tendency scales ^a 1 (certainly not)– 7(<i>certainly</i>)	Trait scales
	To what extent does this person look...
1. Does this person look like someone capable of lifting and manipulating very heavy objects?	1 (weak)–7 (strong)
2. Imagine that this person is one of your flatmates. Does he look like someone who would comply with your wishes about furniture arrangement, decoration, sharing in cleaning, and/or other norms that you suggested?	1 (submissive)–7 (dominating)
3. Does this person look like someone capable of explaining a very difficult theory to you with clarity?	1 (unintelligent)–7 (intelligent)
4. Does this person look like someone who would swallow a hoax? ^b	1 (naive)–7 (shrewd)
6. Does this person look like someone who would turn a cold shoulder to your attempts at a friendly conversation?	1 (cold)–7 (warm)
7. Does this person look like someone who would falsify his student record?	1 (honest)–7 (dishonest)
5. Does this person look like someone who normally says things that are not true?	1 (truthful)–7 (deceitful)
8. Does this person look like someone who would accuse an innocent person of having committed a sexual abuse offense?	
	How does this person look?
	1 (very unattractive), 2 (rather unattractive), 3 (somewhat unattractive), 4 (intermediate), 5 (somewhat attractive), 6 (rather attractive), 7 (very attractive).

^aNumbers indicate the order of the items. Notice that the two deception-relevant items were not placed together (they occupied the 5th and the 8th positions in the questionnaire). This was done to minimize any carryover effect from one question to the other. The order of the items in the trait section of the questionnaire was the same as in this table.

^bMcArthur and Apatow's (1983–84) behavioural-tendency item for naivety–shrewdness asked whether the stimulus-person looked like someone who would believe a far-fetched story told on April Fools' Day. The question in the present experiment may look different, but it is not really. The Spanish word that was used for hoax was *inocentada*, which means either a hoax made on, or a "far-fetched story" told on, Holy Innocents' Day (28th December), the Spanish equivalent to the US Fools' Day (1st April).

^cMcArthur and Apatow (1983–84) asked respondents to rate whether the stimulus-person looked like someone who would look at a classmate's paper for answers during a final exam. Since the present authors' informal observations suggest that many undergraduates might have copied during a final exam, the decision was made of asking about a more serious action with far-reaching consequences.

TABLE 3
Univariate tests of the effect of facial babyishness on the behavioral-tendency and trait scales, and mean ratings

	<i>F</i>	<i>p</i>	η^2	<i>Significance of the Linear trend</i>	<i>Significance of the Quadratic trend</i>	<i>Means</i>		
						<i>BFced</i>	<i>Interm</i>	<i>MFced</i>
Behavioural-tendency scales								
Dishonesty	3.20	.042	.02	.018	.417	3.09 ^a	3.30 ^{ab}	3.64 ^b
Deceitfulness	3.16	.044	.02	.088	.074	3.35 ^a	3.72 ^b	3.69 ^{ab}
Sexual abuse	0.53	.592	.00	.571	.403	3.14 ^a	2.89 ^a	3.11 ^a
Strength	7.13	.001	.05	.000	.743	3.77 ^a	4.09 ^b	4.41 ^b
Submissiveness	1.34	.264	.01	.108	.823	3.86 ^a	3.71 ^a	3.51 ^a
Intelligence	2.04	.132	.01	.210	.107	3.05 ^{ab}	2.96 ^a	3.23 ^b
Naivety	3.49	.032	.02	.011	.497	4.71 ^a	4.63 ^{ab}	4.17 ^b
Coldness	4.12	.017	.03	.005	.675	3.81 ^a	3.97 ^{ab}	4.50 ^b
Trait scales								
Dishonesty	2.65	.072	.02	.035	.405	3.15 ^a	3.36 ^{ab}	3.61 ^b
Deceitfulness	1.46	.233	.01	.133	.392	3.46 ^a	3.36 ^a	3.77 ^a
Strength	13.93	.000	.08	.000	.126	3.73 ^a	4.35 ^b	4.59 ^b
Dominance	15.84	.000	.09	.000	.339	3.59 ^a	4.18 ^b	4.63 ^b
Intelligence	1.06	.347	.01	.247	.401	4.12 ^a	4.08 ^a	3.89 ^a
Shrewdness	9.81	.000	.06	.000	.494	3.72 ^a	4.23 ^b	4.51 ^b
Warmth	6.13	.002	.04	.001	.702	3.60 ^a	3.48 ^a	2.83 ^b

The scales were bipolar (e.g., weakness–strength). In this table higher ratings reflect that perceivers' impressions were closer to the endpoint indicated by the label in the first column, and therefore farther from its opposite. For example, higher ratings in the "strength" row indicate more strength and, consequently, less weakness.

BFced: Babyfaced; Interm: Intermediate; MFced: Mature-faced.

Polynomial and simple contrasts were performed to analyse the trends and the differences between the faces of different levels of facial babyishness.

Means with different superscripts differ significantly.

TABLE 4
Univariate tests of the effect of age on the behavioural-tendency and trait scales, and mean ratings

	<i>F</i>	<i>p</i>	η^2	<i>Significance of the Linear trend</i>	<i>Significance of the Quadratic trend</i>	<i>Means</i>		
						<i>Child</i>	<i>Adult</i>	<i>Old</i>
Behavioural-tendency scales								
Dishonesty	2.58	.077	.02	.463	.031	3.21 ^{ab}	3.53 ^a	3.29 ^b
Deceitfulness	4.29	.015	.03	.022	.081	3.85 ^a	3.41 ^b	3.52 ^b
Sexual abuse	4.24	.015	.03	.016	.125	2.71 ^a	3.17 ^b	3.26 ^b
Strength	198.92	.000	.57	.000	.000	2.18 ^a	4.96 ^b	5.14 ^b
Submissiveness	5.08	.007	.03	.356	.002	3.64 ^a	4.04 ^b	3.39 ^a
Intelligence	16.61	.000	.10	.000	.000	2.51 ^a	3.51 ^b	3.21 ^b
Naivety	13.76	.000	.08	.000	.356	5.08 ^a	4.41 ^b	4.00 ^b
Coldness	11.27	.000	.07	.000	.258	3.48 ^a	4.21 ^b	4.60 ^b
Trait scales								
Dishonesty	0.92	.398	.01	.727	.063	3.40 ^a	3.20 ^a	3.52 ^a
Deceitfulness	2.36	.097	.02	.193	.288	3.68 ^a	3.40 ^a	3.51 ^a
Strength	85.18	.000	.35	.000	.010	3.01 ^a	4.46 ^b	5.20 ^c
Dominance	40.67	.000	.21	.000	.468	3.23 ^a	4.19 ^b	4.98 ^c
Intelligence	1.31	.271	.01	.613	.128	4.20 ^a	3.95 ^a	3.93 ^a
Shrewdness	8.92	.000	.05	.000	.078	3.88 ^a	3.97 ^a	4.60 ^b
Warmth	7.07	.001	.04	.001	.160	3.81 ^a	3.19 ^b	2.91 ^b

The scales were bipolar (e.g., weakness–strength). In this table higher ratings reflect that perceivers' impressions were closer to the endpoint indicated by the label in the first column, and therefore farther from its opposite. For example, higher ratings in the "strength" row indicate more strength and, consequently, less weakness.

Polynomial and simple contrasts were performed to analyse the trends and the differences between the faces of different age.

Means with different superscripts differ significantly.

Multivariate effects

The MANCOVA conducted on the behavioural-tendency scale ratings yielded a significant multivariate effect for the attractiveness covariate, *Wilks'* $\lambda = .91$, $F(8, 295) = 3.56$, $p = .001$, $\eta^2 = .09$. The main effects of age, *Wilks'* $\lambda = .37$, $F(16, 590) = 23.54$, $p < .001$, $\eta^2 = .39$, and facial babyishness, *Wilks'* $\lambda = .89$, $F(16, 590) = 2.29$, $p = .003$, $\eta^2 = .06$, were also significant. The age x babyishness interaction was not significant, *Wilks'* $\lambda = .89$, $F(32, 1089.501) = 1.05$, $p = .390$, $\eta^2 = .03$.

The MANCOVA conducted on the trait scale ratings revealed significant multivariate effects for the attractiveness covariate, *Wilks'* $\lambda = .83$, $F(7, 305) = 8.94$, $p < .001$, $\eta^2 = .17$, the age independent variable, *Wilks'* $\lambda = .57$, $F(14, 610) = 14.31$, $p < .001$, $\eta^2 = .25$, and facial babyishness, *Wilks'* $\lambda = .83$, $F(14, 610) = 4.24$, $p < .001$, $\eta^2 = .09$. In this case the age x babyishness interaction was significant, *Wilks'* $\lambda = .87$, $F(28, 1101.115) = 1.60$, $p = .026$, $\eta^2 = .04$.

Univariate effects

As reported above, physical attractiveness had significant multivariate effects on both behavioural-tendency and trait ratings. Univariate analyses on the behavioural-tendency measures revealed that higher attractiveness ratings yielded higher ratings of honesty, $F(1, 302) = 20.87$, $p < .001$, $\eta^2 = .07$, warmth, $F(1, 302) = 4.59$, $p = .033$, $\eta^2 = .02$, and, to a lesser degree, intelligence, $F(1, 302) = 3.71$, $p = .055$, $\eta^2 = .01$. They also produced a stronger belief that the target-individual would not accuse an innocent of having committed a sexual abuse offense, $F(1, 302) = 7.50$, $p = .007$, $\eta^2 = .02$. Likewise, on the trait scales, higher attractiveness ratings were related to higher ratings of honesty, $F(1, 311) = 14.45$, $p < .001$, $\eta^2 = .04$, truthfulness, $F(1, 311) = 11.37$, $p = .001$, $\eta^2 = .04$, warmth, $F(1, 311) = 21.58$, $p < .001$, $\eta^2 = .07$, and intelligence, $F(1, 311) = 32.21$, $p < .001$, $\eta^2 = .09$. This effect also approached significance for the dominance measure, $F(1, 311) = 3.74$, $p = .054$, $\eta^2 = .01$. Effect sizes were small, however.

Honesty, truthfulness, and sexual abuse

As shown in Table 3, on the behavioural-tendency scale, the higher the facial maturity, the more

dishonesty was attributed to the stimulus-person, $F(2, 302) = 3.20$, $p = .042$, $\eta^2 = .02$; for the linear trend, $p = .018$. A marginally significant effect pointing in the same direction emerged for the trait scale, $F(2, 311) = 2.65$, $p = .072$, $\eta^2 = .02$; for the linear trend, $p = .035$. Table 4 reveals that the stimulus-person's age had only a marginal effect on the behavioural-tendency scale ratings of honesty, $F(2, 302) = 2.58$, $p = .077$, $\eta^2 = .02$. Adults were perceived as the most inclined to commit a dishonest act, for the quadratic trend, $p = .031$. The effect was not significant at all on the trait scale ratings, $F(2, 311) < 1$.

On the behavioural-tendency scales, babyfaced individuals were judged as less inclined to lie than the other people⁵, $F(2, 302) = 3.16$, $p = .044$, $\eta^2 = .02$. However, the effect of babyishness was negligible on the trait scale ratings of truthfulness or deceitfulness, $F(2, 311) = 1.46$, $p = .233$, $\eta^2 = .01$ (Table 3). Similar results were found concerning the age variable (see Table 4). On the behavioural-tendency scale ratings, children were perceived as more inclined to lie than adults and old persons, $F(2, 302) = 4.29$, $p = .015$, $\eta^2 = .03$; the linear trend was significant, $p = .022$. However, the effect of age was not significant on the trait scale ratings, $F(2, 311) = 2.36$, $p = .097$, $\eta^2 = .02$ (Table 4).

As expected, age had a significant effect on the false sexual abuse allegation ratings, $F(2, 302) = 4.24$, $p = .015$, $\eta^2 = .03$. There was a linear ($p = .016$) increase in the perceived likelihood of the stimulus-person falsely accusing somebody of having sexually abused someone else as the stimulus-person's age increased (see Table 4). The hypothesis specifically predicted significant differences between children and nonchildren. This planned contrast (contrast weights were 1, $-.50$, $-.50$) was significant, $p = .004$, as was the one comparing children with faces that were intermediate in age, $p = .016$, and old faces, $p = .010$; but the younger and older adults' faces did not differ significantly on this measure, $p = .171$ (Table 4). In sum: Children were perceived as less inclined than adults and aged individuals to accuse an innocent of having committed a sexual abuse offence. The effect of facial babyishness on this measure was insignificant, $F < 1$.

No significant interaction between age and babyishness emerged for honesty–dishonesty, truthfulness–deceitfulness, or sexual abuse ratings.

⁵As shown in Table 3 neither the linear, $p = .088$, nor the quadratic trend, $p = .074$, reached statistical significance. Additional contrasts revealed that babyish and mature faces differed slightly, $p = .088$, and that babyish and intermediate faces differed substantially, $p = .016$. No significant difference was apparent between intermediate and mature faces, $p = .477$. A Helmert contrast was then run, and it indicated that deceitfulness ratings of the babyfaced photographs were significantly lower than those of the other two faces taken together, $p = .016$.

Other perceptions

Table 3 reveals that babyfacedness had significant effects on three of the remaining behavioural-tendency scale ratings: strength, naivety, and warmth. Babyishness also significantly influenced perceptions of strength, dominance, shrewdness, and warmth on the trait scales. Although not all faces differed significantly, for all the measures the linear trend was significant, and the direction of the effects was as predicted (see Table 3). In addition, as shown in Table 4, the stimulus-person's age significantly influenced all the remaining measures not described in the previous section on the behavioural-tendency scales, and all but intelligence on the trait scales. Although not all faces differed significantly, the linear trend was significant for all the measures but submissiveness on the behavioural-tendency scale (though polynomial contrasts indicated that strength ratings on both scales, and intelligence on the behavioural-tendency scale, also fitted a quadratic trend; see also means and comparisons in pairs in Table 4). Except for submissiveness on the behavioural-tendency scale, all the significant age effects were in the predicted direction (see Table 4).

The age \times babyishness interaction was only significant for the trait scales measuring perceptions of strength, $F(4, 311) = 4.09$, $p = .003$, $\eta^2 = .05$, and dominance, $F(4, 311) = 2.90$, $p = .022$, $\eta^2 = .04$. It also approached significance for the behavioural-tendency ratings of strength, $F(4, 302) = 2.34$, $p = .055$, $\eta^2 = .03$. In all cases, the increases in ratings caused by the babyishness variable were more pronounced for adult faces than for those of extreme ages (child's and old person's).

DISCUSSION

Except for the intelligence ratings (which, as expected, were not affected by facial babyishness) and honesty ratings (which were not under the influence of the stimulus-person's age), the experimental manipulations significantly affected all the dependent variables on at least one of the two scales used to measure them. Babyish faces were perceived as more honest, more truthful, weaker, more submissive, more naïve, and warmer than the mature faces. This supports the babyface overgeneralization effect (Montepare & Zebrowitz, 1998; Zebrowitz, 1997), according to which those individuals whose faces merely resemble those of babies are attributed baby-like traits. It also suggests that the results of facial

babyishness research conducted in the US can confidently be generalized to South-European countries. Also, certain improvements of the present study, such as using realistic computer-manipulated faces and a between-subjects design, strengthen McArthur and Apatow's (1983–84) conclusions. Age also produced significant linear increases in perceptions of the stimulus-person's truthfulness, tendency to falsely accuse someone else of having committed a sexual abuse offense, strength, dominance, shrewdness, and coldness.

These effects were not caused by the target's attractiveness. However, in accordance with psychosocial research that has shown that attractiveness has a strong influence on a number of social judgments (e.g., Langlois et al., 2000), in this experiment higher ratings of perceived attractiveness were associated with higher attributions of positive qualities such as warmth, intelligence, honesty, and sincerity. This is consistent with the attractiveness-halo effect (e.g., Langlois, 1986) and the what-is-beautiful-is-good stereotype (Dion et al., 1972).

Honesty, truthfulness, and sexual abuse

In general, the predictions concerning honesty and truthfulness were supported by the data. The more the facial maturity of the stimulus-person, the more dishonest that person was perceived on the behavioural-tendency scale. A marginal trend in the same direction was apparent on the trait scale. These results seem consistent with those of the published literature conducted in the US (see, e.g., Montepare & Zebrowitz, 1998). However, as argued in the Introduction, among Spaniards honesty (*honestidad*) is not the same as truthfulness, but is closer to uprightness. Therefore, the present results are actually consistent with those indicating that babyfaced individuals are seen as more righteous or good than mature-faced individuals (Montepare & Zebrowitz, 1998). The age of the stimulus person's had no effect on how honest that person was considered to be. At best, a marginal effect that fitted a quadratic trend was found on the behavioural-tendency scale.

Although, unexpectedly, neither babyishness nor age had any significant effect on how truthful or deceitful the stimulus-person was rated on the trait scale, these variables were significant for the behavioural-tendency ratings. In line with the prediction, babyish faces were judged as the least deceitful. This finding is consistent with the notion that an individual's facial maturity might

be responsible for the demeanour bias that has been found in deception-detection research. However, it is important to stress that the participants in the present experiment did not have to judge the veracity of statements purportedly produced by more or less babyfaced persons. Instead, they had to express their impressions of these individuals. Ongoing research using the same faces as in the present study is examining whether facial maturity has an influence on judgments of credibility in a deception-detection task.

In the behavioural-tendency scale children were judged as more mendacious than adult and old people. This is consistent with the belief that children are imaginative and unreliable witnesses. However, in accordance with the hypothesis, respondents considered that children would be less inclined than adults to lie about a sexual abuse allegation. The belief that children never lie about sexual abuse, as well as social concern with regard to sexual offences against children (e.g., Sachsenmaier & Watson, 1998), are probably the factors that lie behind this finding.

Perceptions of older faces

Simple contrasts revealed that few differences between adult and old faces reached significance (see superscripts in Table 4). In addition, these differences did not support the notion that negative stereotypes toward aged targets were in operation. Even for competence ratings such as strength and shrewdness, older stimulus-persons were rated more positively than younger adults on the trait scales. Older faces were also rated as more dominant than the younger adult ones on both scales. These results might be due to our failure in choosing a sufficiently aged face (see footnote 2). Also, extant research has found that more negative attitudes are apparent toward the so-called old-old (aged 70 and over) than toward the young-old (65–70) (Youmans, 1977). Therefore, research into the social perception of old people should use targets from both of these old-age categories.

Two significant babyishness by age interactions indicated that the increases in strength and dominance on the trait scale measures caused by facial maturity were more pronounced for adult faces than for those that were very young or old. Probably, there were age indicators in the extreme faces whose influence on these ratings was stronger than that of the babyishness cues that were manipulated. However, babyishness and age did not interact significantly for any other

dependent measure. This indicates that, in general, the effects of facial babyishness were the same across all the age range represented in this experiment. It also indicates that the effects of age were the same across all the facial babyishness levels.

Unexpected results

Results on the behavioural-tendency scales measuring submissiveness departed from what was expected. First, although the means were in the right direction, the effect of facial babyishness on this measure did not reach statistical significance. More importantly, the effect of age was not linear, but quadratic: Those judged as the most submissive were not the children, but the adults. An unfortunate selection of the dependent variable may account for this unexpected finding. Respondents were asked to indicate the extent to which the stimulus-person would comply with certain domestic norms. It is possible that respondents considered that children tend to be disobedient and undisciplined, so that they would hardly comply with these norms.

Finally, although children were perceived as being less capable than adults of explaining a very difficult theory with clarity, age did not have any significant effect on the trait scale ratings of intelligence. This does not necessarily question the notion that children are considered to be less intelligent than adults. There is no reason to take an adult's intelligence as a reference in judging how intelligent the child in the photograph is. A child's photograph is probably judged in reference to other children or to an ideal, internalized notion of the average child, not in reference to an adult. In this regard, the behavioural-tendency measure of intelligence was different, however. Respondents might be aware that a difficult theory is something alien to a young child (*any* young child), who could hardly understand it and would probably be unable to explain it to someone else.

Conclusions and implications

The present study replicated the findings of facial babyishness research in a South-European country, using sophisticated computer technology to manipulate real faces and a conservative, between-subjects design. Thus, apparently, Europeans also hold facial stereotypes rooted in overgeneralizations of the adaptive responses to babies. It is however important to note that,

although in the present study most effects were statistically significant, effect sizes were in general very small.

A number of studies have shown that facial babyishness influences not only the way people are perceived, but also the way they are treated in a number of contexts (Montepare & Zebrowitz, 1998). The present study has the merit of having focused specifically on the influence of facial maturity on impressions of truthfulness, as a separate construct, different from the more general notion of uprightness that is reflected in the Spanish word *honestidad* (honesty). In the behavioural-tendency scales, babyishness influenced not only honesty (uprightness), but also truthfulness. It is therefore possible that, merely because some people have a certain facial appearance, they are considered to be truthful, whereas others are considered to be deceptive. This might have important consequences in legal contexts, in particular if, as suggested by the present results, in addition to being untruthful, mature-faced defendants are also considered not to be righteous.

Another finding of the present study merits attention: Children were considered to be more deceptive than older individuals, unless the lie concerned a sexual abuse offence, in which case they were rated as the most truthful. In view of the increase in child sexual abuse allegations over the last decades (e.g., Lamb, 1994), and the number of unconfirmed cases in divorce or custody and visitation disputes (e.g., Thoennes & Tjaden, 1990; Wakefield & Underwager, 1991), this result is of significance. Judgments based on first impressions are not necessarily accurate, and it is worrisome that children are given more or less credibility depending on the topic. In view of the important consequences of misjudging children's allegations in sexual abuse cases, precise credibility-assessment tools are needed (see Manzanero, 2001).

Manuscript received June 2003

Revised manuscript received September 2003

REFERENCES

Alonso-Quecuty, M. L. (1996). Las teorías implícitas existentes sobre los niños testigos como moduladoras de la aceptación o rechazo de su testimonio [Implicit theories on child witnesses as modulators of the acceptance or rejection of their testimony]. In F. del Castillo & J. D. Valdivieso (Eds.), *Congreso de psicología jurídica de Castilla y León. Libro de actas* (pp. 157–158). Salamanca: Colegio Oficial de Psicólogos de Castilla y León.

- Atzwanger, K., & Grammer, K. (1993). *Babyness and sexual selection*. Poster presented at the 23rd International Ethological conference?Torremolinos, Spain.
- Baartman, H. E. M. (1992). The credibility of children as witnesses and the social denial of incestuous abuse of children. In F. Losel, D. Bender, & T. Bliesener (Eds.), *Psychology and law: International perspectives* (pp. 345–351). Berlin: Walter de Gruyter.
- Berry, D. S. (1990). Taking people at face value: Evidence for the kernel of truth hypothesis. *Social Cognition*, 8, 343–361.
- Berry, D. S. (1991). Attractive faces are not all created equal: Joint effects of facial babyishness and attractiveness on social perception. *Personality and Social Psychology Bulletin*, 17, 523–531.
- Berry, D. S., & Brownlow, S. (1989). Were the physiognomists right? Personality correlates of facial babyishness. *Personality and Social Psychology Bulletin*, 15, 266–279.
- Berry, D. S., & McArthur, L. A. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, 48, 312–323.
- Berry, D. S., & McArthur, L. A. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, 100, 3–18.
- Berscheid, E., & Walster, E. (1974). Physical attractiveness. *Advances in Experimental Social Psychology*, 7, 157–215.
- Bond, C. F. Jr., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. *Journal of Experimental Social Psychology*, 21, 331–345.
- Brigham, J. C., & Spier, S. A. (1992). Opinions held by professionals who work with child witnesses. In H. Dent & R. Flin (Eds.), *Children as witnesses* (pp. 93–111). Chichester, UK: Wiley.
- Brownlow, S. (1992). Seeing is believing: Facial appearance, credibility, and attitude change. *Journal of Nonverbal Behavior*, 16, 101–115.
- Brownlow, S., & Zebrowitz, L. A. (1990). Facial appearance, gender, and credibility in television commercials. *Journal of Nonverbal Behavior*, 14, 51–60.
- Bruce, V., & Young, A. (1998). *In the eye of the beholder. The science of face perception*. Oxford: Oxford University Press.
- Bull, R. (Ed.) (2001). *Children and the law: The essential readings*. Malden, ME: Blackwell.
- Butler, R. N. (1969). Age-ism: Another form of bigotry. *The Gerontologist*, 9, 243–246.
- Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review. *Psychological Bulletin*, 113, 403–439.
- Damhorst, M. L. (1990). In search of a common thread: Classification of information communicated through dress. *Clothing and Textiles Research Journal*, 8, 1–12.
- Dent, H., & Flin, R. (Eds.) (1992). *Children as witnesses*. Chichester, UK: Wiley.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Clairton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74–118.

- Dion, K. K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285–290.
- Enlow, D. H. (1982). *Handbook of facial growth* (2nd ed.). Philadelphia, PA: Saunders.
- Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, 111, 304–341.
- Friedman, H., & Zebrowitz, L. A. (1992). The contribution of typical sex differences in facial maturity to sex role stereotypes. *Personality and Social Psychology Bulletin*, 18, 430–438.
- Goodman, G. S., & Bottoms, B. L. (1993). Child victims, child witnesses: *Understanding and improving testimony*. New York: Guilford Press.
- Kite, M. E., & Johnson, B. T. (1988). Attitudes toward older and younger adults: A meta-analysis. *Psychology and Aging*, 3, 233–244.
- Lamb, M. E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse and Neglect*, 18, 1021–1028.
- Langlois, J. H. (1986). From the eye of the beholder to behavioral reality. In C. P. Herman, M. P. Zanna, & E. T. Higgins (Eds.), *Physical appearance, stigma, and social behavior: The Ontario symposium*, Vol. 3 (pp. 23–51). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390–423.
- Manzanero, A. L. (2001). Procedimientos de evaluación de la credibilidad de las declaraciones de menores víctimas de agresiones sexuales [Procedures for assessing the credibility of children victims of sexual abuse]. *Psicopatología Clínica, Legal y Forense*, 1, 51–71.
- Masip, J., Garrido, E., & Herrero, C. (2002). La detección del engaño sobre la base de sus correlatos conductuales: La precisión de los juicios [The nonverbal approach to the detection of deception: Judgmental accuracy]. *Anuario de Psicología Jurídica*, 12, 37–55.
- Masip, J., Garrido, E., & Herrero, C. (2003). When did you conclude she was lying? The impact of the moment the decision about the sender's veracity is made and the sender's facial appearance on police officers' credibility judgments. *Journal of Credibility Assessment and Witness Psychology*, 4, 1–36.
- McArthur, L. Z., & Apatow, K. (1983–84). Impressions of baby-faced adults. *Social Cognition*, 2, 315–342.
- McArthur, L. Z., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology*, 18, 165–192.
- Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. *Advances in Experimental Social Psychology*, 30, 93–161.
- Montepare, J. M., & Zebrowitz, L. A. (2002). A social-developmental view of ageism. In T. Nelson (Ed.), *Ageism: Stereotyping and prejudice against older persons* (pp. 77–125). Cambridge, MA: MIT Press.
- Nelson, T. (Ed.) (2002). *Ageism: Stereotyping and prejudice against older persons*. Cambridge, MA: MIT Press.
- Sachsenmaier, S. J., & Watson III, H. A. (1998). Do psychological techniques for investigating child sexual abuse allegations meet the requirements of scientific evidence and the admissibility of expert witness testimony? Unpublished manuscript.
- Thoennes, N., & Tjaden, P. G. (1990). The extent, nature, and validity of sexual abuse allegations in custody/visitation disputes. *Child Abuse and Neglect*, 14, 151–163.
- Todd, J. T., & Mark, L. S. (1981). Issues related to the prediction of craniofacial growth. *American Journal of Orthodontics*, 79, 63–80.
- Todd, J. T., Mark, L. S., Shaw, R. E., & Pittenger, J. B. (1980). The perception of human growth. *Scientific American*, 242, 132–144.
- Wakefield, H., & Underwager, R. (1991). Sexual abuse allegations in divorce and custody disputes. *Behavioral Sciences and the Law*, 9, 451–468.
- Youmans, E. G. (1977). Attitudes: Young-old and old-old. *The Gerontologist*, 17, 175–178.
- Zebrowitz, L. A. (1997). Reading faces. *Window to the soul?* Boulder, CO: Westview Press.
- Zebrowitz, L. A., Fellous, J. M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as over-generalized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review*, 7, 194–215.
- Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced individuals across the life span. *Developmental Psychology*, 28, 1143–1152.
- Zebrowitz, L. A., Montepare, J. M., & Lee, H. K. (1993). They don't all look alike: Individual impressions of other racial groups. *Journal of Personality and Social Psychology*, 65, 85–101.
- Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, 22, 1258–1269.
- Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrace, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology*, 15, 378–396.
- Zuckerman, M., Koestner, R., & Driver, R. (1981). Beliefs about cues associated with deception. *Journal of Nonverbal Behavior*, 6, 105–114.