CHILDREN'S JUDGEMENTS OF SAMENESS-DIFFERENCE AMONG SCHEMATIC FACES

Carolyn Roy

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

1983

CHILDREN'S JUDGMENTS OF SAMENESS-DIFFERENCE AMONG SCHEMATIC FACES

by

CAROLYN ROY

Dissertation submitted towards the degree of Ph.D. in Psychology
at the University of St. Andrews, 1982.

ABSTRACT

Researchers have found that young children often make "confusion errors" in judging visual stimuli for sameness. That is, children are reported to treat as "just the same" stimuli that are not identical in all respects. Explanations of these confusion errors have focussed on separate aspects of children's performance: language development (e.g., acquisition of relational terms), or perceptual and cognitive development (e.g., visual scanning or selection of criterial features). The latter appear more successful in explaining the *types* of errors manifested. For instance, Taylor (1973) found that children, judging sameness among schematic faces, consistently correctly matched particular features and confused others, and later evidence (Wales, pers. comm., 1974) suggested their responses may be affected by the presence or absence of a visual frame around the face.

The present series of experiments investigated potential influences on judgments of sameness among sets of schematic faces by children of pre-school and early school age. Variables examined included presence of a visual frame of reference around the face, type of stimulus array, task requirements, and salience of stimulus features relative to each other.

Initial results confirmed that, in tasks like matching-from-sample or pair-comparison, children systematically matched only certain features of the faces and confused other features. Type of task (including stimulus and response variables) and presence/absence of a visual frame interacted with each other to influence the response patterns. Subsequent experiments suggested that neither objective visual salience of one stimulus feature over another, nor selection of a visual criterial attribute, satisfactorily accounted for the observed response patterns. A more likely contributor was subjective weighting of features on grounds of their contextually-afforded significance. Faces were not compared feature-by-feature; instead, the criterion seemed to be sameness of affective expression conveyed.

It is argued that many of the confusion errors were not due simply to linguistic or perceptual immaturity. Rather, in circumstances where adults take "same" to refer to identity across all features, the children took it as "same kind of thing": their responses reflected their bases for <u>classifying</u> the faces (i.e., by affective content) rather than ability to judge identity. Having judged unidentical <u>faces</u> as the same, a considerable number of children indicated in justifying their judgments that while certain <u>features</u> were the same, others were not. Thus they appeared able to understand and use "same" in its <u>various</u> adult senses, but interpreted the task instructions differently from adults.

CERTIFICATE

I hereby certify that Carolyn Ann Mary Roy has completed nine terms of research under supervision after being admitted as a research student under Ordinance General No. 12, 1967. She has fulfilled the conditions of the resolution of the University Court, 1967, No. 1, and is qualified to submit the accompanying dissertation in application for the degree of Doctor of Philosophy.

Research Supervisor

December, 1982

DECLARATION

I hereby declare that the accompanying dissertation was written solely by me, that the research reported therein was conducted by me alone, and that it has not been submitted in part or entirety in any previous application for a higher degree.

Carolyn A.M. Roy

December, 1982

ACKNOWLEDGEMENTS

This research and its documentation would not have been completed without the assistance of a number of people, who made invaluable contributions both in terms of criticism and ideas and on the more practical side of providing children as subjects and producing the accompanying report.

My thanks are due to my three research supervisors. Roger Wales guided me at the start of this work, helping me to choose a course for research and providing background material and stimulating discussion during its early phase. When he left St. Andrews for a post in Australia, Bob Grieve took over; his comments on the completed first experiments and ideas for the design of stimuli and experiments in the next phase were valuable. When he, too, left for Australia, Frank Quinault stepped in to supervise the final parts of the research and the writing of the dissertation. His suggestions -- and encouragement -- while I was planning and writing the report, and his critical reading of the drafts, were greatly appreciated.

I gratefully acknowledge the cooperation of the staff and assistants in the various centres in which I gathered the experimental data. Not only did they willingly allow me to disrupt their schedules; all were also most helpful in providing information on subjects' ages, etc., and space for me to conduct the experiments, in encouraging children to take part, and in fending off inquisitive but non-participating children. The centres involved were: Greyfriars Primary School, Kinnessburn Playgroup, Puffin Playgroup, the St. Andrews Nursery Classes, the St. Leonard's Playgroups (all in St. Andrews); Leuchars Primary School; and the Cowley St. John Playgroups (Oxford). Additionally, the Superintendent of Albany Park Residence of the University of St. Andrews kindly gave me permission to advertise for subjects among families staying there during the vacation. A few further children of acquaintances were tested privately at home. Adult responses to the stimulus figures and in the experimental tasks were obtained from students in Albany Park and Deans Court, St. Andrews. My thanks go to all of these adults and children. Their enthusiasm was infectious.

Laura Gutman assisted in typing some of the final manuscript, with speed and care. I am particularly grateful to her for coping with the lists in Appendix 3 and with the tables throughout the report.

During the initial years of this research programme, I was supported by a Research Studentship from the Medical Research Council of Great Britain.

My parents have also provided financial support, as well as constant encouragement throughout my studies.

CONTENTS

APPENDICES

PART 1

CHILDREN'S CONFUSION ERRORS IN JUDGMENTS OF SAMENESS

Judgments of sameness and the related operations of discrimination and classification are routine human activities. Infants begin early on to sort their world of people and objects into distinct individuals and categories, more so as they start to acquire names for them. As Eve Clark recently commented,[1] fundamental to such categorisation is the capacity to decide when things are similar enough, or too dissimilar, to be treated as the same.

From the outset, then, children need some grasp of sameness and difference. By the age of three or four, their appreciation of these concepts appears well-developed, at least in everyday situations. They readily solve puzzles where they have to spot the differences between two not-quite-identical drawings, and can successfully play the kinds of picture matching and sorting games that are common fodder in preschool centres. Yet, when confronted with such tasks in psychological experiments, those young children are shown to be far from adult in their treatment of sameness and difference. When asked to find in a test array an item that is <u>different</u> from the target indicated by the experimenter, they may choose one that is <u>identical</u>. Or they may give a false "same" response, treating <u>unidentical</u> stimuli as <u>just the same</u>.

The occurrence of false "same" judgments, or confusion errors, has been demonstrated over and over in a variety of experimental settings. It appears that the young child often treats items as "the same" when they are merely <u>similar in some respect(s)</u>, and in circumstances where an adult would call them the same only if they were <u>identical in all</u>

respects. The phenomenon of children's confusion errors is the main concern of the research programme to be reported here.

Gregson (1975, pp. 15-18) distinguishes two types of false "same" judgment. In one, the subject simply fails to detect any difference between unidentical items. In the other, he does perceive a difference but nonetheless decides the items are the same. Proposed explanations for children's confusion errors reflect this dichotomy. Errors that are believed to be of the first type attract perceptual explanations -- the child failed to see the difference. Accounts based on language are associated with the second type of error -- the young child may let perceived differences pass because he does not understand the word "same" to mean completely identical; he thinks partial similarity suffices. In both these cases, the child is seen as failing as a result of his (perceptual or linguistic) immaturity. There are, in addition, accounts that lean more towards cognitive functioning, implicating strategies of information-processing and decision-making. Here, the child is regarded as thinking in ways that are not necessarily immature as such, but are not the ways that an adult would follow given the same situation.

To evaluate theory in the light of empirical findings in this field is a daunting task. Besides direct examination of same-different decisions, evidence can be drawn from studies of classification, discrimination-learning, recognition, concept-attainment, and so forth, which involve judgments of identity or similarity and difference to some degree. The methodological diversity, together with the variety of stimuli and of types of stimulus transformations employed, make it difficult to collate and compare results across the numerous experiments that have been conducted. A selective outline of some of the reported studies and the main findings and theories will be presented to provide a general background to the programme of new research to be described.

Despite the range of task and stimulus variables employed, certain regularities have emerged in the data on children's confusion errors. One of the most interesting of these, for present purposes, is that confusion errors occur in rather systematic patterns. Some sorts of discriminanda are consistently frequently confused, even by children as old as seven or eight years, while others are almost invariably correctly distinguished by three-year-olds. For instance, as we shall see, differences in orientation of a figure are often confused by preschool children; but they rarely confuse different shapes.

In the meantime, let us take for granted the occurrence of such systematic biasses in the distribution of confusion errors. For, this finding that errors are associated with only certain types of stimulus transformations allows us immediately to comment on the language-based view that asserts that young children have only partial understanding of the meaning of the word "same". The plan will then be as follows.

After noting some problems for that language-oriented approach, we shall return to examine more closely what sorts of discriminanda are confused and in what circumstances, and how those errors might otherwise be explained. Perceptual and cognitive factors that have been forwarded to account for the findings include children's visual strategies for inspecting and comparing stimulus figures, their modes of perceiving and interpreting stimuli, and their decision-rules or -criteria. Much of the evidence points to the child's increasing differentiation of stimulus properties with age. Studies of concept- and category-formation and recent extensions of research on separability versus integrality of of stimulus dimensions will be noted in this connection. Part 1 concludes with a brief comment on psychometric approaches to similarity judgments.

The rest of this report is devoted to the new experimental investigations of the present programme, which was initiated in the wake of some intriguing previous findings. Throughout, the stimuli consist of sets of schematic drawings of faces. Part 2 (Experiments 1 - 3) examines the effects of type of task (manipulating stimulus array and the kind of response required of the child) and of presence/absence of an external visual frame surrounding the stimulus (a perceptual cue) on the degree to which children confuse the outer shape or, instead, the internal detail of the faces.

The remaining experiments concentrate on the differential distribution of confusion errors among only the <u>internal</u> face features. In Part 3 (Experiments 4 - 6), the relative contributions of perceptual and cognitive factors are assessed. Are confusions due to failure to notice a feature or detect a difference (visual), or to deliberate ignoring of it (cognitive)? The possible roles of physical saliency of stimulus features, subjects' visual scanning strategies, and subjects' assignment of differential priorities among stimulus features, are considered. Part 4 continues to address those issues, examining the subjects' justifications for their sameness judgments during Experiments 2 to 6. Implications as to the children's decision-criteria and understanding of the experimenter's test question are discussed. Experiment 7 employs a method intended to clarify and restrict for the child the possible interpretations of the task requirements. The question is whether, when the strictness of the decision-rule expected by the experimenter is emphasised in the test instructions, the child will then adopt the adult's criteria for sameness.

Part 5 (Experiment 8) constitutes a final word on the differential distribution of confusion errors among the internal face features. Do such schematic faces have a "criterial attribute" that children will consistently treat as more important than other features, or are they judged on some more global basis? The test is whether or not children will redeploy their attention among the attributes when their variation alters the global character of the faces. A summary of the findings and general discussion concludes the report.

Language-oriented explanations for confusion errors

The sort of account meant here is that which is founded in the "semantic-feature theory" of child language acquisition. It stems largely from the work of Eve Clark (e.g., 1971, 1973a,b), who extended the linguists' method of componential analysis to children's acquisition of lexical items. It has been applied especially to the acquisition of temporal and spatial prepositions, deictic verbs and demonstratives, comparative expressions, etc.* What concerns us here is its application in the domain of so-called polar adjective pairs (e.g., big-wee, more-less), for the terms same-different have been included in this group (see, e.g., E.V. Clark, 1973a; H. Clark, 1970; Donaldson & Wales, 1970; Wales *et al*., 1976; Webb *et al*., 1974).

----

* For a general picture of this approach to language acquisition, see (e.g.) Donaldson & Balfour (1968), Campbell & Wales (1970), Wales & Campbell (1970), Seymour (1974), Macrae (1976), Gathercole (1978), as well as the E. Clark papers cited above. Clark & Clark (1977, Chs. 11 & 13) provide a useful overview. Wales *et al*. (1976), Grieve & Hoogenraad (1976) and Macrae (1978), among others, offer contrary evidence, while Richards (1979) gives an extensive critique of Eve Clark's approach.

This is not the place for an extensive critique of semantic-feature theory. Suffice it to say that this approach in general is not without problems. The terms same-different are especially troublesome. Put simplistically, the reasoning within this framework goes along the following lines.

Sameness-difference is viewed as a dimension. For adults, the term "same" is held to apply to the positive pole of the dimension, referring to perfect identity. The young child initially acquires the idea that both "same" and "different" have to do with the general domain of (dis)similarity. Gradually, he attaches further semantic features to the word "same" that allow him to refine his understanding of its meaning and to locate it towards one end of the continuum. At this point in his development, he takes "same" to mean "similar in some way". The addition of features and refinement continue until "same" comes to be located at the pole. Now the child takes it to refer to complete identity in all respects. Meanwhile, the child at first equates "different" with "same", realising only that both words have to do with similarity. Once he has grasped that "same" belongs at the positive end of the continuum, he begins to distance "different" from it, attaching the negative so that it becomes "not identical".

At the theoretical level, one problem with this treatment of same-different is that they do not behave like the other polar adjective pairs. "Same", for example, is regarded as restricted to the extreme end of the dimension (perfect identity), while "different" refers to any point not at the positive pole. Other polar adjective pairs do not exhibit this extreme skew. There is, moreover, no single dimension of sameness-difference in reality. Objects may be simultaneously the same with respect to one quality, such as size, and different in another, such as colour. Judgments of sameness may, indeed, involve perceptual-cognitive processes that are independent from those operating in judgments of difference (e.g., Tversky, 1977).

At the empirical level, such a treatment does neatly encompass the developmental side of children's sameness-difference judgments. The gradual decrease with age in number of errors made is accounted for by the child's increasing refinement of the semantics of "same" and "different" (e.g., Donaldson & Wales, 1970; Webb et al., 1974 -- but see also Glucksberg et al., 1976, for possible criticism of their work). But there is one important aspect of the error data that this approach fails to deal with. This is the repeated finding, already noted, that certain sorts of discriminanda are more confusable than are others. If it were simply the case that the young child understood "same" to mean merely "similar in some way", presumably any kind of similarity would do. His confusion errors would be randomly distributed across the range of stimulus transformations studied. In practice, however, the errors fall in quite systematically biassed patterns. So, while accounting for the developmental decrease in number of errors, semantic-feature theory fails to explain the types of confusions that occur. At best, then, this language-oriented explanation is anadequate on its own.

Let us now return to consider exactly what types of confusion errors occur and alternative explanatory hypotheses.

Confusion errors and type of stimulus transformation

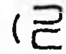
Among the principal sorts of stimulus transformations to have been studied are variations of form, orientation, colour, size, detail, etc. Vurpillot (1976) provides an extensive review of experimental and theoretical work on children's judgments of sameness-difference in relation to these stimulus variables. A sketch of the main trends in the findings and proposed explanations is provided below.

<u>Orientation and reversal</u>

Numerous experiments have demonstrated children's frequent failure to distinguish between stimuli that differ by orientation or reversal. Even within this general domain, however, some sorts of variations seem more confusable than others (see Vurpillot, 1976, Ch. 3). A 45^{0} rotation such that one stimulus is horizontal and the other vertical (e.g., | vs —) elicits hardly any confusion errors from preschool children. Diagonal lines of different orientations (\ /) are more often confused with each other. Rotations of 180^{0} (e.g., d vs p) and inversions (b P) are moderately confusable; while reversals that turn the figures into "mirror images" of each other (b d) produce the most numerous -- and, developmentally speaking, the most persistent -- confusion errors.

Considerable research effort has been spent on children's suscepti-bility to this last kind of reversal confusion. One commonly-cited study is that of Rudel and Teuber (1963), who demonstrated preschool children's failure to distinguish such reversals with pairs of single oblique lines or of horseshoe shapes (⊏ ⊐). Using a kind of habitua-tion and transfer paradigm, Bornstein and Gross (1978) found that babies made no distinction between reversed stimulus pairs consisting of faces, of single lines, or of horseshoe shapes, although they were sensitive to other differences in orientation with those stimuli.

Following Rudel and Teuber's study, several investigators demon-strated the importance of the <u>alignment</u> of reversed stimuli relative to each other. Huttenlocher (1967a,b), for instance, found left-right reversals of horseshoe shapes presented in left-right alignment (⊏ ⊐)

to be highly confusable; but they were generally successfully discriminated when they were presented in up-down alignment (). Similarly, up-down reversals were confused far more often when displayed in up-down alignment () than in left-right alignment (). Sekuler and Rosenblith (1964) obtained alignment effects with rotated triangles and circles containing a gap in the circumference; and Vurpillot (1976) reported these effects with novel geometric forms (e.g., ). Apparently, such reversals are confused when they are presented as true mirror images of each other, but can be differentiated when not aligned as exact mirror images.

Explanations for these mirror-image confusions have implicated neurological mechanisms (to do with symmetry and asymmetry of representations between the cerebral hemispheres), perceptual factors (to do with visual scanning paths and spatial relationships) and cognitive-social factors (to do with the adaptiveness of ignoring certain orientation differences in the real world). Bryant (1974), Vurpillot (1976) and Bornstein and Gross (1978) review them. The current trend favours the last-mentioned of the three schools of thought over the more traditional "mirror-image" explanations. There are, however, several irregularities in the data that no theory as yet seems able to deal with single-handedly.

One is that up-down mirror images are generally confused less often than are left-right ones, even when stimulus alignment is controlled. That is, children do not simply confuse <u>any</u> mirror-image pair. It is hard to explain this in terms of a general problem with the visual comparison of stimuli in this kind of spatial arrangement or of a general tendency deliberately to ignore mirror-image transformation differences.

Another is that the type of task and the type of response required of the child can affect whether or not he distinguishes the stimuli. Jeffrey (1958), Hendrickson and Muehl (1962) and Rosenblith (1965) have shown that preschool children could differentiate mirror-image pairs when they signalled their decision by a motor response, such as pushing a button or pulling a lever on the left or on the right, where they failed with a verbal response. Differences have been found between visual discrimination and copying performance on the same set of stimuli, although pretraining may help to reduce errors in both types of task (e.g., Jeffrey, 1966; Huttenlocher, 1967b; Olson, 1970; Strayer & Ames, 1972). Over and Over (1967) demonstrated that, while six-year-olds made mirror-image confusions of single oblique lines when the stimuli were presented in pairs as in the conventional simultaneous comparison or successive discrimination-learning methods, even four-year-olds could perform successfully when those stimuli were presented in a matching-from-sample format, where the subject has to choose from a multiple array one figure that matches a "standard" given by the experimenter. Bryant (1974) also emphasises the importance of such methodological factors. Mirror-image confusions are apparently not a simple perceptual matter, then.

Furthermore, despite the common report of these confusion errors, children -- and even babies -- have been shown to be generally sensitive in other respects to differences of orientation (e.g., Rice, 1930; Watson, 1966; Olson & Baker, 1969; Bornstein & Gross, 1978). It is likely that the type of stimulus figure -- its complexity and inherent degree of directionality -- is important. Certain kinds of stimuli, especially those depicting familiar items (such as faces), tend to be treated as mono-oriented. Rotations from the canonical orientation are perceived as altering the nature of such a figure, and are correctly discriminated. Thus Sinclair and Piaget (1968) found that two squares, one in canonical orientation and one rotated, were judged to be different entities by four-year-olds; not until the age of six did children begin to realise that, regardless of orientation, both figures were the same square form.

Poly-oriented stimuli -- those lacking any canonical orientation -- are another matter, particularly if they are unfamiliar forms as well. Williams et al. (1976) obtained an increase with age in orientation confusions in pair-comparison of eight-sided random polygons. Even adults may sometimes exhibit such orientation effects as a function of type of stimulus. Gibson and Robinson (1935) found that adult subjects recognised outlines of familiar, mono-oriented stimuli much better when they were presented in canonical orientation than when they were rotated, while poly-oriented figures were recognised equally well in all orientations.

These sorts of results might suggest the influence of experience and of real-world knowledge about when orientation is likely to be important or not. Lila Ghent Braine, in contrast, has emphasised the visual side of children's performance on such orientation differences. She conducted a series of investigations using stimuli that varied in whether or not they were interpretable as mono-oriented and in degree of directionality, as specified by, e.g., the presence of a gap or some detail at one end of the figure (Ghent, 1960, 1961; Ghent et al., 1960; Ghent & Bernstein, 1961; Braine, 1965; Barroso & Braine, 1974). Her results led her to conclude that under-fives differ from older children in terms of which part of a stimulus figure they visually fixate first, and in which direction visual scanning of the figure proceeds. (These results are discussed in more detail in Part 3, pp. 133-134.)

Wohlwill and Wiener (1964) agreed with Ghent Braine that while young children are sensitive to orientation, their perceptual judgments may be helped or hampered by the nature of the stimuli in terms of degree of directionality, location within the figure of a "focal area", and location of distinguishing features. This kind of research can be criticised, however, for making assumptions about what defines canonical orientation, directionality, focal area, etc., and for generally "going beyond the information given" (see Vurpillot, 1976). In any event, as we have already noted, such a visual scanning-strategy account alone fails to handle the complexity of the findings from the experiments of other investigators.

Some final remarks on the orientation-confusion research. With stimuli that vary in multiple ways, orientation is typically the property that is most often confused. Taylor and Wales (1970) gave children matching-from-sample tasks with geometric figures that varied in number of sides, degree of curvature, degree of openness, etc., as well as in orientation. Differences of shape and what Taylor and Wales called "discrete" transformations, such as openness versus closure, were readily discriminated by three-year-olds. So-called "continuous" variations, as in number of sides or size of gap, were associated with some confusion errors. Orientation differences ("continuous") were still confused by five-year-olds. Those authors assumed that the children were differentially weighting the various sorts of transformation, taking some kinds of difference to be less important than others, i.e., deliberately disregarding orientation and giving precedence to shape.

McGurk (1972) argued a similar case, but on sturdier grounds. He used an assortment of stimuli, some abstract and some realistic.

Children were given two tasks. In the first, where orientation alone was varied, the subjects showed good identification of orientation differences. Subsequently, they received a matching-from-sample task in which size and/or colour varied as well as orientation. Now, orientation was confused while size and colour were usually correctly matched. McGurk concluded that children can discriminate orientation, but when other stimulus properties also vary, orientation becomes relatively less salient. If the nature of the task or the stimuli is such as to direct the child's attention to the orientation differences, confusions are avoided.

In short, taking the Taylor-Wales and McGurk results along with the earlier-mentioned (p. 11) findings of orientation effects in older children and adults (Williams et al., 1976; Gibson & Robinson, 1935), it appears that orientation and reversal confusions are better explained in terms of cognitive-social factors than purely perceptual ones.

What about other sorts of confusion errors? These can be dealt with more briefly, since the results are usually less ambiguous in interpretation than are those on orientation and reversal. Vernon (1966) and Vurpillot (1976) provide useful overviews.

Shape and its complexity

If orientation is the stimulus property that attracts the greatest number of confusion errors, the opposite is true for shape (contour, form, etc.). Only shapes that are both highly complex and highly alike fail to be discriminated by children over three years old; then there is a gradual increase with age in the amount of complexity the child can handle. Santa Barbara and Paré (1965), for example, gave children stimulus figures composed of star shapes varying in type of star and in number of prongs, for pair-comparison. Three-year-olds could successfully distinguish the different stars as long as they had only a few prongs, but confusion errors increased as the number of prongs increased. Five-year-olds, on the other hand, could differentiate the more complex stars. Brown and Goldstein (1967) obtained a similar result with random polygons: five- and six-year-olds quickly distinguished the simpler ones, but made errors as the number of sides increased to 12 or more. Errors in such situations are likely due more to the younger child's limited capacity for information-processing than to anything else.*

Shape, size and colour

When shape is varied along with other stimulus properties such as size or colour, it is generally the case that shape dominates as the basis for sameness judgments. Long (1941) demonstrated that four- to seven-year-olds confused size more often than shape but that they were able to learn size discriminations quite readily. Farkas and Elkind (1974) gave children of five to nine years the task of comparing geometric figures for size, at three distances. Confusion errors were fewer (1) with age, (2) with reduced distance, and (3) when the standard (larger) figure was on the left than on the right of the test figure. This implicates the visual side of the comparison process more than usual.

Whereas McGurk (1972) found that preschool children discriminated size and colour differences equally well in matching-from-sample, Oléron (1962) found that they confused colour more often than height, which in turn was confused more often than form, in a pair-comparison task. (By age seven, subjects distinguished variations in all three properties

---

* For the possible influence of other sorts of complexity, see (e.g.) Zusne (1965), Halford & Macdonald (1977).

with equal success.) The reason for the discrepancy between those two sets of results is not clear. Perhaps it lies in the differing methodologies employed, or perhaps the relative physical saliencies of size and colour differed in the two studies.

Colour and shape combinations have been manipulated in a number of experiments. There is general agreement that responding on the basis of form increases with age, while responding on the basis of colour decreases; but there is disagreement as to which of the two, form or colour, actually predominates for younger children (age five or less). Again, these discrepancies might arise from methodological or physical saliency differences across experiments. Reichard et al. (1944) found that in a free classification task four- and six-year-olds consistently sorted by form rather than colour while five-year-olds did the reverse. In constrained classification (where the experimenter stipulates the basis of grouping), however, those children tended to sort randomly; only after age six could subjects consistently classify by the named attribute and ignore the other, irrelevant one. Corah (1966) in a three-choice-oddity matching task, and Suchman and Trabasso (1966) and Odom and Mumbauer (1971) in concept-learning tasks, all found that shape dominated over colour at age five and colour responses decreased further with age. Wales et al. (1976) found that shape was preferred to colour when the stimuli were solid, realistic objects (books, pens, matchboxes, covered in coloured plastic), while colour was preferred when the shapes were less "meaningful" (two-dimensional drawings of geometric forms). (See also Campbell et al., 1976, for an interesting longitudinal study.)

A study by Hale and Lipps (1974) directly addressed the issue of whether the reason that an attribute is apparently overlooked in such tasks is because it is not perceived or because it is assigned low priority. In a three-choice-oddity task, children over about age five chose matches of shape and confused colour, while the reverse held for children under five. A subsequent component-selection task demonstrated that, whichever attribute the subject had used in the three-choice-oddity format, he was nonetheless aware of differences in the other, previously-confused attribute and could successfully differentiate both shape and colour when required to. This suggests that deliberate selective inattention, rather than perceptual failures, may underly confusions of an attribute when the stimuli vary in multiple ways (cf. McGurk's [1972] findings, mentioned earlier on pp. 12-13).

Internal detail

Eliane Vurpillot and her colleagues have systematically investigated the possible role of perceptual failures in confusion errors. In a series of experiments, children were given pictures of houses for pair-comparison (e.g., Vurpillot, 1968; Vurpillot & Moal, 1970; Berthoud & Vurpillot, 1970; Vurpillot *et al*., 1971). In the typical format, each house had six windows; houses varied in whether the contents of the windows differed, or whether the same windows were present but in different orders (e.g., such that the two houses were mirror images of each other). Houses differing in window-contents generally attracted fewer false "same" judgments than did houses differing only in window-permutation. Children who made confusion errors were found to inspect the pictures incompletely, to scan in unsystematic or inappropriate paths, to fixate few of the relevant details, and to inspect the stimuli for a shorter time than did children who made successful pair-comparison discriminations.*

---

* Other investigators whose work implicates such visual processes include Zaporozhets (1965) on the degree to which children's eye movements systematically follow contour lines; Faw and Nunally (1968) and Mackworth and Bruner (1970) on children's visual fixations of parts of a stimulus as a function of "high information value", high contrast, novelty, pleasantness,

While emphasising the role of immature visual inspection strategies, Vurpillot nevertheless admits that some confusion errors may have more to do with linguistic and cognitive factors such as the young child's unrefined notion of sameness or his assignment of differential importance among stimulus attributes. This was evident in a two-part study she conducted using pairs of complex pictures, such as schematised landscapes, as stimuli (Vurpillot, 1969). The two pictures in each pair differed by only minor alterations among the mass of identical detail. Differences were made up from among four types of transformation: (1) a detail was present in one picture, absent in the other; (2) a detail was halved in size in one picture; (3) a detail was rotated 45° or was displaced to a new location (classed as a single kind of variation); (4) the shape of a detail was transformed (but, Vurpillot claimed, not so that it became a different kind of object -- e.g., a moon might change from full to crescent).

In the first half of the study, children were given a detection task. They were told that there were differences between the two pictures, and that they should point them out. Four-year-olds readily spotted differences involving shape and presence/absence of a detail. Size differences were not usually indicated until age seven, while displacement/rotation shifts were often missed even by seven-year-olds. Detail-changes that had the effect of altering a major contour line were more often indicated than were those same transformations if they were performed on an interior detail.

/continued/ etc. Elkind (1964) and Elkind <u>et al.</u> (1964) comment on children's unadultlike perception of ambiguous figures and of parts versus wholes, respectively.

In the second experiment, the task was conventional pair-comparison for same-different judgment. The order of confusion errors paralleled the order already observed in the detection study: displacement/rotation and size transformations were frequently confused. Pair-comparison performance, however, was generally inferior to detection performance; and children who gave false "same" pair-comparison judgments could still point to differences between the pictures.* While failure to detect a difference probably contributed to the observed distribution of confusion errors with these highly complex pictures, then, there was undoubtedly more to the errors than that. The children appeared to be deliberately ignoring some of the differences in the pair-comparison task.

A study by Ricciuti (1963 -- reported in Gibson, 1969, p. 346) also points to the role of this kind of selective inattention. Subjects aged three and upwards were given a matching-from-sample task with stimuli that varied in shape (geometric forms) and in details such as the presence/absence of dots, external protrusions and indents in the contour line. The children chose matches based on gross similarity of contour and confused differing details, even though they demonstrated that they could perceptually distinguish the details. Such findings reinforce a point made by Imai and Garner (1965): there is an important distinction to be made between what a subject <u>can</u> do and what he <u>does</u> do. Those authors, too, found that whereas physical detectability governed the pattern of responses in one kind of task, constrained classification, subjective preferences for one stimulus property over another, regardless of physical discernibility, dominated under free classification of the same stimuli -- and their subjects were adults.

---

* Again, cf. the similar method and findings of McGurk's (1972) two-part study; see pp. 12-13.

From generalisation to differentiation: 1. Criterial attributes

Immature strategies of visual inspection of stimuli, as we have seen, may contribute to children's confusion errors but are unable on their own to account for all of the findings. In particular, they fail to explain why some kinds of differences are confused when it is likely or certain that the children can and do detect them. Children are sensitive to differences in orientation, size, colour, etc. As Vurpillot (1976) remarks, children can correctly name and match colours long before they will successfully distinguish them in a pair-comparison task. And several investigators have reported that following a confusion error the child will proceed to point out ways in which the stimuli he just judged to be the same are in fact different. Oléron (1962), for instance, quotes one of his French subjects who confused size as saying, "Si, ils sont tous pareils, mais il y en a un plus petit et un plus grand" (see also, e.g., Taylor & Wales, 1970; Vurpillot & Moal, 1970; Kemler & Smith, 1979). Is it then that children employ different decision-strategies or looser criteria of sameness than adults?

In the 1960s, Eleanor Gibson and her colleagues conducted a series of experiments in the discrimination-learning tradition that gave rise to an influential alternative theory of children's confusion errors (e.g., Gibson et al., 1962; Gibson, 1963, 1965, 1969, 1971; Pick, 1965; see also Schaller & Harris, 1974). A set of basic "letter-like forms" was constructed to serve as standards. Transformations were applied to produce for each standard a series of variants. The child's task was to achieve discrimination of each variant from its standard. In keeping with other experimenters' findings on confusion errors, transformations that effected "discrete" alterations of contour, such as break versus closure, were readily discriminated by the youngest subjects (four-year-olds). "Continuous" contour changes, such as substitution of a curved line for a straight one, and transformation by rotation or reversal, provoked a fair number of errors from four-year-olds but performance improved markedly with age. Slanting the figure in various directions to produce a transformation of perspective yielded the hardest variant for the children to distinguish. The error rate here was near 80% for subjects up to age six, and improved only slowly thereafter (to 60% for eight-year-olds).

Gibson's account of these error patterns focussed on "criterial attributes" or "distinctive features".* The argument runs as follows. Young children have only limited cognitive capacity for information-processing and storage. The child adapts to this by encoding stimulus information in ways that abbreviate and/or omit certain aspects of it. Through experience, he develops a notion of what sorts of information are likely to be important and what can normally be safely disregarded in given situations. In discriminations and sameness judgments, he therefore treats some stimulus attributes and some kinds of transformations as more <u>critical</u>, or <u>criterial</u>, than others, and ignores differences that he regards as "minor".

What is most criterial for differentiation of letters of the alphabet are topological differences such as closure of a gap, presence/absence of a line, etc. (compare H vs A, P vs R) that noticeably alter the shape of the outline; perspective is normally unimportant. The observed order of discriminability among the transformations of the experimental letter-like

---

* The term "criterial attribute" is here preferred, since "distinctive feature" can be ambiguous. Some authors use the latter to refer to any feature that differs between stimuli instead of to the one, of several possible stimulus properties, that subjects treat as criterial to sameness-difference.

forms illustrates this. The argument is extendable to the patterning of confusion errors with other sorts of discriminanda. The form and function of objects are supposed to be of special interest to the preschool child. He therefore tends to assign higher priority to shape than to, say, size or colour, and lowest to differences in orientation and perspective which he has found in general less likely to be important variables.

What develops with age is then not so much the child's ability to make perceptual discriminations as his cognitive capacity to encode larger amounts of information efficiently. Hand in hand with that growth goes the refinement and modification of the child's decision-criteria in making sameness-difference judgments.

This explanation of how the child proceeds from apparent overgeneralisation to finer and finer differentiation of stimulus properties fits reasonably well with much of the empirical data on confusion errors. It also ties in rather neatly with work in other fields that has indirect bearing on sameness-difference judgments.

Related areas: categories and concepts

The classification of items into categories is obviously relevant. (For major studies in this area in general, see, e.g., Vygotsky, 1962; Inhelder & Piaget, 1964; Bruner *et al.*, 1966.) A common finding has been that young children, if they sort consistently at all, group items on the basis of global functional or perceptual similarities;* around age

---

* There has been some debate as to precisely what are the stimulus characteristics used by young children in categorical and conceptual behaviour. While the Clarks have taken a perceptual criterial features kind of viewpoint, Katherine Nelson has argued the primacy of more global functional bases (see, e.g., E. Clark, 1973b; H. Clark, 1970, 1973; Nelson, 1972, 1973, 1974, 1978). In practice, since perceptual and functional characteristics are often confounded in stimuli, results can be ambiguous. (See also Saltz *et al.*, 1972.)

six or so, they attend increasingly to separate stimulus attributes; finally, they progress to abstract bases of sorting. Studies whose results are in keeping with this pattern include those of Sigel (1953), McConnell (1964), Rossi (1964), Ricciuti (1965), Al-Issa (1969), Pick and Frankel (1974). Rock et al. (1972) showed that adult subjects resorted to global bases of similarity and overlooked differences of detail (unless specifically directed what to look for) when a recognition task was made difficult by the use of highly complex figures and tachistoscopic presentation.

The work of Eleanor Rosch and her colleagues on "natural categories" ties in with the general view that children proceed with age from undifferentiated to differentiated and ultimately abstract modes of encoding stimuli (e.g., Rosch, 1973a,b, 1977; Rosch & Mervis, 1975; Rosch et al., 1976). Some aspects of this research also have interesting implications with regard to specific findings in the confusion-error literature.

In Rosch (1973a), it is suggested that certain kinds of geometric figures (particularly those of good Gestalt form) may act as "core" or "prototypical" instances of categories, while transformations of those figures are fuzzier, "boundary" category members. A perfect square, for instance, is a prototypical tetragon; a variant such as a trapezoid, while still belonging to the tetragon category, is not prototypical (and so on for the circle, triangle, etc., categories). Rosch (1973a) demonstrated that American children and Dani adults (the latter having no experience of geometric forms), given the task of learning names for instances of concepts, did appear to treat such stimuli in terms of prototypes and sets of variants, some of which were closer to the prototype than others. To take the example of the squareness concept, a variant containing one curved side was rapidly learned, i.e., was seen as close to the prototypical square; while an irregular tetragon, which subjects took longer to associate with the "square" name, was treated as a boundary instance and dissimilar from the prototypical square.

In her 1975 paper, Rosch went on to propose that category prototypes serve as "cognitive reference points", and that non-prototypical members are perceived and judged in relation to those reference points rather than to each other. In one experiment, adult subjects were required to indicate perceived psychological distances among stimuli by positioning them in physical space. One of the stimulus sets consisted of single straight lines of various orientations. Rosch originally proposed that within the domain of orientation there are three reference points: the strictly vertical, the strictly horizontal, and the 45^0 diagonal. In fact, subjects used only the vertical and horizontal axes for reference. Diagonal lines were perceived only in relation to those two axes. This suggests that horizontality and verticality are distinct categories of orientation and that diagonals may be boundary members of those categories instead of separate concepts in their own right.

There are obvious links between this and the confusion-error research. On the basis of Rosch's studies, one would predict that prototypes of distinct categories should not be confused with each other. The data confirm this: perfect squares are readily distinguished from perfect circles; horizontal and vertical lines are not confused. Secondly, transformations that only slightly mar goodness of fit to the category core should be confusable with the prototype, whereas transformations that turn the variant into a poor category member should be distinguishable from the prototype (cf. children's actual performance on perspective versus topological transformations of a standard). Boundary instances may, however, be confusable with each other (cf. the observed orientation confusions between differing diagonals, as opposed to correct discrimination between diagonal and vertical, or diagonal and horizontal).

The notion of prototypes has also attracted attention in the field of artificial-concept attainment. While this research does bear on the issues of children's ability to distinguish stimulus attributes, weighting of one attribute over another, bases of similarity and difference judgments, etc., the area is too complex and too indirectly relevant for detailed consideration here. Bolton (1972) provides a fairly comprehensive review of the field.* Authors who have given specific consideration to the abstraction of prototypes (sometimes called "schemas") include Bruner et al. (1956), Pick (1965), Posner (1969), Reed (1972), Lasky (1974), Posnansky and Neumann (1976). Developmentally speaking, the findings tend to agree with those of the classification and discrimination studies: Preschool children perceive the stimuli globally; by age six or so, they differentiate relevant stimulus attributes, but if they do engage in prototypical representation of stimulus information their prototypes are incomplete -- they abstract only part of the necessary information; after about age eight they perform competently, attend to all relevant attributes, switch basis of responding when required to, and abstract informative prototypes.**

---

* For a general picture of the range of stimuli employed and of theories forwarded, see also (e.g.) Bruner et al. (1956), Bourne & Restle (1959), Kendler & Kendler (1962), Huttenlocher (1964), McConnell (1964), Suppes & Rosenthal-Hill (1968), Steinberg (1974), Bozinou & Goulet (1974).

** Performance in tasks like classification, discrimination-learning, concept-attainment, etc., may be influenced by factors such as the meaningfulness of stimuli, their amenability to verbal labelling, imagery, etc. Since results in these areas are, however, contradictory and inconclusive, they will not be discussed here. For illustration, see Vanderplas & Garvin (1959), Kendler & Kendler (1962), McConnell (1964), Rossi (1964), Clark (1965), Faw & Nunally (1968), Danks (1970), Mwanalushi (1974), Saltz & Finkelstein (1974), Wales et al. (1976), Jörg & Hörmann (1977).

<u>From generalisation to differentiation: 2. Integral-separable stimulus dimensions</u>

While the majority view, then, is that preschool children focus on general similarity among stimuli and come gradually with age to discriminate attributes individually, a few researchers have obtained results that apparently run totally counter to the main trend. Several of the studies by Saltz and his colleagues have produced this reversed pattern (e.g., Saltz & Sigel, 1967; Saltz <u>et al</u>., 1972 -- Kogan, 1976, cites further examples). As early as 1919, Claparède concluded that children learn first about difference. Claparède questioned children about similarity (Are a bee and a wasp the same? How? Which are most alike: a bee and a bird or a bee and a flower? Etc.); the children, even five-year-olds, responded by indicating differences and rarely mentioned resemblances.

Although relatively few and far between, such counter-findings are potentially troublesome for the Gibsonian generalisation-to-discrimination theory. Kogan (1976) suggests a way of reconciling the discrepancies: they have to do with the nature of the stimuli. Young children follow the more usual overgeneralisation pattern (resulting in confusion errors in same-different judgments) when stimulus differences are located in <u>details</u> and thorough visual inspection is required for them to be perceived. They distinguish between stimuli, however, when the differences are immediately noticeable from overall Gestalt. Saltz and Sigel's (1967) stimuli were photographs of boys' faces, varying in identity of boy and in facial expression. Saltz <u>et al</u>.'s (1972) stimuli were realistic pictures of various categories of entities (assorted food and furniture items, kinds of animals, etc.). In both these cases, differences were readily apprehensible from Gestalt properties. Hence Saltz's finding of apparent "overdiscrimination" in five-year-olds. According to Kogan, those subjects were reacting to global dissimilarities rather than discriminating detail.

Other investigators who have used faces as stimuli have sometimes observed this kind of pattern too (e.g., Odom & Lemond, 1974). Carey and Diamond (1977) and Diamond and Carey (1977) gave children photographed faces wearing extraneous paraphernalia -- hat, spectacles, distinctive hair-style, earrings, etc. -- for matching-from-(serially-presented)-sample. The task was to identify the individual's face, regardless of paraphernalia. Five-year-olds performed correctly when the faces were of people <u>familiar</u> to them; here, their judgments did seem to rely on global facial configuration. With unfamiliar faces, however, even eight-year-olds mismatched. They were apparently distracted by the paraphernalia. Those authors interpreted the confusions of unfamiliar faces to indicate that the children were not perceiving the stimuli configurally, but were attending to the salient detail. There is an alternative, Gibsonian interpretation, however: the children might have been choosing matches on the basis of the overall similarity afforded by the configuration of face-<u>plus</u>-paraphernalia. Those studies are therefore inconclusive in terms of the generalisation/differentiation issue.

Firmer evidence -- and an elaboration of Kogan's (1976) explanation -- is available, however, in the research of Garner and some recent extensions of it. Working mainly with adult subjects, Garner in the 1960s addressed such topics as Gestalt goodness versus complexity of pattern and subjects' use of stimulus properties versus abstracted rules for partitioning stimulus sets into subsets in encoding information (e.g., Garner, 1962, 1966; Garner & Clement, 1963; Whitman & Garner, 1963). Garner (like Gibson, e.g., 1971) emphasised that people perceive and process stimulus information in ways designed to reduce uncertainty (see also Osler & Kofsky, 1965).

More recently, Garner has concentrated on one aspect of the relationship between stimulus structure and subjects' modes of perceiving that is pertinent to the question of overgeneralisation/differentiation (e.g., Garner, 1970, 1974; Garner & Felfoldy, 1970). According to these authors, some kinds of stimuli are composed of dimensions that can be independently varied and independently perceived -- size and colour, for example. These are called <u>separable</u> dimensions, and a subject who does differentiate them is perceiving in the separable mode. With other stimuli, however, variation in one dimension brings about a change in the appearance of the whole stimulus, as if affecting the other dimensions. Thus variation of saturation alone, or of brightness alone, is perceived as an overall change of colour. Such dimensions are called <u>integral</u>. Integral dimensions are normally not perceptually differentiable. But it is not always the case that separable dimensions <u>are</u> perceptually differentiated. Structurally separable stimuli may sometimes be perceived integrally (see further Treisman & Gelade, 1980).

This model has now been applied to children's perceptions and judgments of stimuli. In a series of experiments varying integrality and separability of stimulus dimensions, Deborah Kemler and Linda Smith have demonstrated that young children perceive integrally stimuli that older subjects (from about age eight upwards) perceive as separable (e.g., Smith & Kemler, 1977, 1978; Kemler & Smith, 1978, 1979; Baron, 1978, summarises their work and cites further references). In tasks such as free or constrained classification and three-choice-oddity matching, five-year-olds therefore based their responses on <u>global</u> similarity (cf. generalisation), while older subjects attended to stimulus properties individually and gave dimensionally-based responses (differentiation). With certain kinds of stimuli and tasks, the two perceptual-judgmental modes yielded the same outcome. In other contexts, however, the two modes produced quite distinct response patterns. When the task called for attention to separate dimensions and the stimuli were separable, older subjects' performance was superior to that of younger subjects (by such measures as speed of sorting). But in other cases, the integral mode of attending to global similarity had a facilitatory effect, so that the younger subjects performed better than the older. (Kagan et al., 1963, made similar observations, although before Garner's model was available.)

The Kemler-Smith research thus manages to knit the apparently anomalous results of Saltz (etc.) into the main Gibsonian framework. That is, children do seem typically to overgeneralise before they come to differentiate. But whereas most experimental methods favour fine discriminatory behaviour, in some (now specifiable) task-stimulus contexts the younger integral, wholistic modes of perceiving and encoding stimuli are actually advantageous.

A note on psychometric studies of sameness-difference

Psychometric approaches to perceptual and conceptual phenomena are being applied increasingly in the area of similarity-dissimilarity. Authors who have devoted attention to classification and similarity relations include Sjöberg (1965), Torgerson (1965), Reed (1972), Tversky (1977), Krumhansl (1978), Schwarz and Tversky (1980) and Lian (1981).*
Gregson's (1975) book contains thorough reviews of empirical, methodological

---

* Among others, Shepard (1964), Shepard & Chipman (1970), and Levelt (1970), while not dealing directly with sameness-difference judgments, comment on relevant aspects of methodology and data analysis.

and theoretical issues in the psychometrics of similarity. These studies have tended to deal as much with the question of appropriate means of data analysis as with the patterns of subjects' responses themselves. What has often been of interest is the degree of fit between the perceived, or subjective, structure of the stimulus set and its objective structure, the latter supposedly specifiable on purely physical grounds independently of the former.

With respect to sameness-difference judgments, there are several problems in the psychometric approach. One, noted by Torgerson (1965), is that a distinction needs to be made between similarity as a "basic" (perceptual) relation and similarity as a "derivative" (cognitive) relation. Both Torgerson and Gregson (1975) comment on the potential complexity of the cognitive processes involved and on the fact that people may switch between various cognitive strategies. Existing psychometric models are inadequate because they are largely tied to the physical, the objectively measurable, and as yet lack the machinery to handle the psychological -- the complexity and subjectivity of subjects' decisions about similarity. Gregson sums up the problem: "... there are many varieties of similarity, some of which may be mathematically well-behaved and others not ..." (1975, p. 2). Since, as we have seen, young children's sameness judgments do not always mirror the objective structure of the stimuli and are possibly even more subjectively-based than are those of adults, psychometric methods would seem inappropriate for the study of confusion errors. False "same" judgments in fact rarely feature in the psychometric literature, partly because the nature of the enterprise precludes this as an issue, and partly because the subjects are usually adults, not children.

One study with adult subjects that has given rise to subsequent research with children (including, indirectly, the series of experiments to be described in the following parts of this report) is Tversky and Krantz's (1969) test of their "interdimensional additivity hypothesis". Their stimulus set consisted of schematic drawings of eight faces, each composed of three distinct features (attributes): shape of head, mouth and eyes. There were two variants, or levels, of each attribute. (Figure 1 of Appendix 1 illustrates a stimulus set constructed along similar principles to that of Tversky and Krantz.)

Tversky and Krantz predicted that the perceived degree of dissimilarity among stimuli would depend partly on the <u>number</u> of actually differing attributes, a quantitative measure. A pair of faces differing in two attributes would be judged less alike than a pair differing in only one attribute. (As in traditional discrimination-learning tasks, the focus here is on difference rather than sameness.) Besides this quantitative feature-count, Tversky and Krantz allowed a qualitative component to judgments of relative dissimilarity among stimuli: subjects might treat one <u>kind</u> of variation as making more of a difference than another kind. They might, for instance, judge a pair of faces differing only in shape of head to be less alike than a pair differing only in eyes. Now, the prediction from the additivity of dimensions hypothesis is that a pair differing in shape <u>and mouth</u> will then be judged less alike than a pair differing in eyes <u>and mouth</u>. That is, the rank order of dissimilarity for pairs differing in two attributes will follow the order for pairs differing in one attribute. Tversky and Krantz's predictions were borne out by their data. Through psychometric analyses, stimulus pairs could be ranked according to degree of perceived dissimilarity, and the order of dissimilarity was found to be preserved when a further dimension of difference was added across stimulus pairs.

How does this work relate to children's sameness-difference judgments? Although Tversky and Krantz did not themselves address the issue of false "same" responses, their framework has been extended to children's confusion errors. John Taylor conducted a wide-ranging investigation of children's sameness-difference performance, reported in his doctoral dissertation (1973). One of Taylor's experiments was stimulated by the Tversky and Krantz study outlined above. It was the data on confusion errors emerging from that experiment, together with apparently conflicting observations by another investigator, that formed the starting point for the research programme described in the remainder of this report.

PART 2

EFFECTS OF PRESENCE/ABSENCE OF VISUAL FRAME AROUND STIMULI AND OF TYPE OF TASK ON DISTRIBUTION OF CONFUSIONS

The Taylor experiment in question is the third reported in his dissertation (1973). In it, he followed Tversky and Krantz (1969) for the design of one of his stimulus sets — eight schematic faces. The task was matching-from-sample. Children (aged 3;10 to 7;6 years) were presented with an array consisting of the eight faces plus, at the foot, a "standard" identical to just one in the array. (Figure 1 of Appendix 1 shows the stimuli and their order in the array.) Each child was asked to "find a picture up here [in the array] that's just the same as this one [the standard]", for all eight standards.

One of Taylor's aims at this point was to determine how stimulus pairs that were incorrectly chosen as matching (i.e., that were confused) would be distributed in both quantitative and qualitative terms -- i.e., with regard to both number and type of attributes confused. On the basis of the Tversky and Krantz work, Taylor made two predictions:

(1) that the frequency of confusion errors would directly parallel the number of actually identical attributes -- e.g., that stimuli sharing two attributes would be more often confused than stimuli sharing just one attribute;

(2) that the "order of similarity" exhibited for matched stimulus pairs that actually differed in one attribute would determine the order for chosen pairs differing in two attributes. (This he called the "ordering hypothesis"; cf. Tversky and Krantz's additivity hypothesis.)

This second prediction rests on the assumption (not initially foregrounded by Taylor) that not all attributes are "equal". It is in fact the <u>attributes</u>, rather than the stimulus pairs, that would exhibit an order-of-similarity effect. If there is indeed any systematic ordering among the particular stimulus pairs (wrongly) chosen as matching, it would simply reflect that the child is overlooking one <u>attribute</u> more than another.

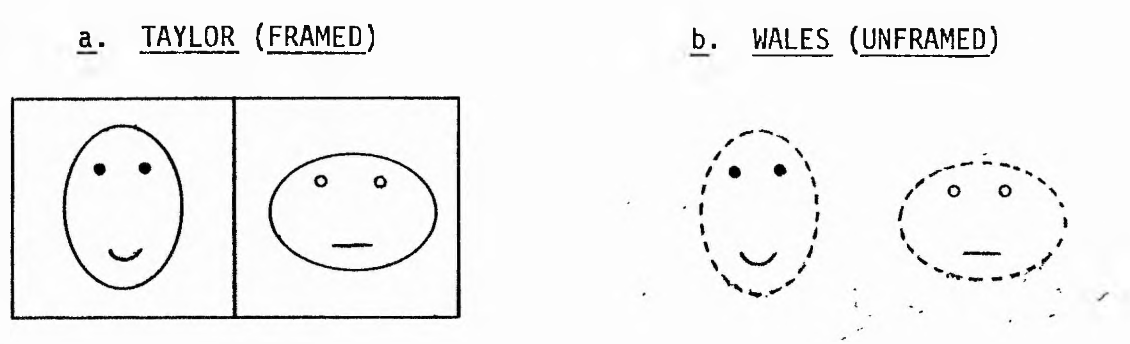
It was Taylor's data as they relate to this assumption, rather than to either of his predictions themselves, that stimulated the present research. The important finding, resulting from both multidimensional scaling and hierarchical clustering analyses of the children's choices as match-to-standard, was that <u>shape of head was the attribute most often correctly matched, while mouth was the most often confused</u>. Despite some variability across subjects, the pattern was essentially unchanged when individual response patterns were considered: of the 48 children whose choices showed a consistent error pattern, two-thirds correctly matched shape of head, and correct matches of eyes tended to be more frequent than were matches of mouth.

What made these data especially interesting was an apparently contradictory finding by Roger Wales (pers. comm., 1974).[2] In a demonstration of children's confusion errors in sameness judgments, Wales used a stimulus set modelled directly on Taylor's (and hence on Tversky and Krantz's, 1969), but observed that the <u>internal features (eyes, mouth) were more often correctly matched than was shape of head</u>.

Whence the contradiction? One salient difference between Wales's demonstration and Taylor's experiment was that, although the discriminanda themselves were the same in the two cases, the visual background of the individual stimuli was not. That is, both employed the same <u>faces</u>, drawn in black ink on white card and having the same three attributes each with their two levels. But in Taylor's stimulus set, not only were the ellipses depicting shape of head drawn in black; each entire face was set within a square, black frame (as shown in Figure 1a, below). Wales, in contrast, cut each face out of the card around the shape of head, so that the ellipse representing shape was given not in black ink but by the cut-out outline; and Wales's stimuli had <u>no</u> surrounding frame external to the head (see Figure 1b). (Notice that the illustrations of the stimuli in Appendix 1 represent Taylor's <u>framed</u> set.)

<u>Figure 1</u>

Illustration of two of the set of eight schematic faces used by Taylor and Wales. Taylor's faces were each surrounded by a square frame; Wales's were cut out around the outline of the head.



(Solid lines indicate what was drawn in black ink; dotted lines indicate an outline that was cut out, not drawn.)

It is easy to imagine several possible reasons why the children should have chosen matches on the basis of shape of head with Taylor's stimuli, and on the basis of internal features with Wales's. It could be, for instance, that the presence of a surrounding frame "draws the eye" of the child towards the edges of the stimulus figure, leading him to overlook the internal features. Conversely, when the frame is absent <u>and</u> the shape of the head is cut out instead of outlined in black, children may have difficulty detecting the shape (because it is less salient than the internal features) or may actively disregard it, deliberately considering only the ink-drawn parts of the figure. Or perhaps Taylor's subjects were using as a cue the visual-spatial relationship between the borders of the square frame and the ellipse of the head shape. That is, in the vertical orientation the ellipse is close to the frame at top and bottom but distant from it at the sides, while the reverse is true for the horizontal ellipses. So the surrounding square may serve literally as a frame of reference, highlighting the visual difference between the two levels of shape of head without affecting discriminability of internal features.

Whatever the reason, however, the Taylor/Wales findings point again to the operation of perceptual-cognitive factors rather than purely linguistic ones in the occurrence of confusion errors. To explore further the perceptual-cognitive aspects of children's errors in sameness judgments, and in particular the effect of presence versus absence of a frame around stimulus figures, became the initial objective, then. Clearly, the first task was to attempt a replication of the Taylor/Wales findings.

EXPERIMENT 1

Matching-from-sample

Taylor's design was followed as closely as possible, from the description of the third experiment in his dissertation (1973).

Materials

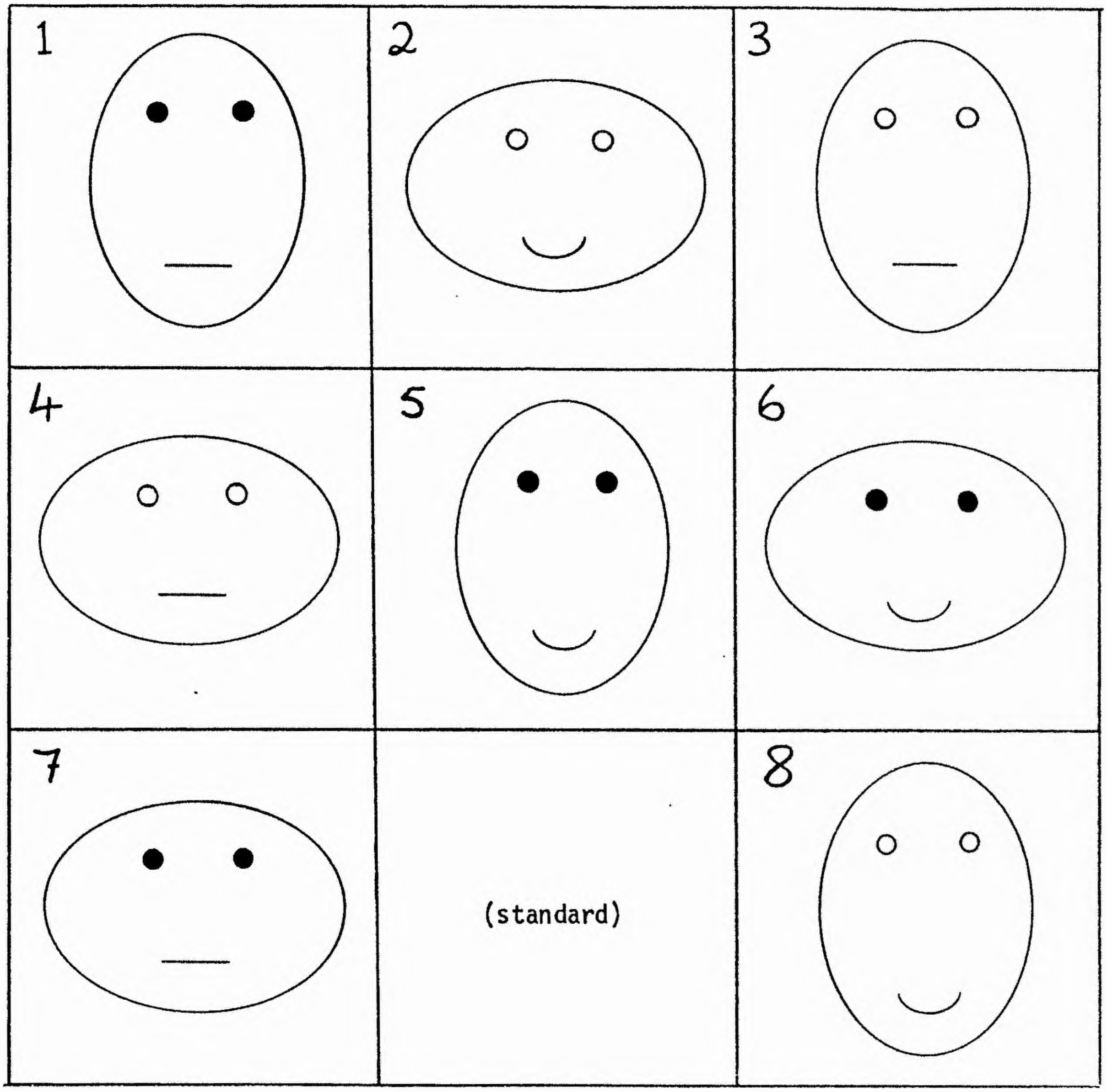
As in the Taylor and Wales work, the stimulus faces were drawn in black ink on white card, and consisted of three attributes each having two variants. These were: Shape (vertical vs horizontal ellipse); Eyes (circles either left white in the middle or filled in black); and Mouth (straight line vs curved-up line). These attributes and levels in all combinations generate a total of eight schematic faces. The set is illustrated in reduced scale in Figure 1 of Appendix 1; for ease of reference, it is also reproduced in the text below (Figure 2, p. 37). Two of the faces, representing between them both levels of each attribute, are shown actual size in Figure 2 of the Appendix. The ellipses depicting Shape measured 9.5 by 6.5 cm across the axes.

Two such sets of eight faces were produced:
The Framed set paralleled Taylor's stimuli. A 33-cm-square white board was divided by straight black lines into nine squares each of side 11 cm. The eight faces were mounted in the squares, leaving the centre bottom frame blank. The outlines of the nine squares on the board thus provided the square frames around the faces (see Figure 2, next page).
The Unframed set copied Wales's stimuli. The eight faces were cut out around the elliptical outline of the head, and were displayed on a matt grey background. The spatial array was as for the Framed set, with the same distance between faces but without the dividing lines that served as frames. Again, the centre bottom space in the 3 x 3 array was left blank.

A further two groups of eight faces were constructed, mounted individually on white cards, to serve as "standards". One of these groups matched the Framed set, each face being surrounded by an 11-cm-square black frame drawn around the edges of the card. The other matched the Unframed set, with each face cut out. All stimuli were covered in washable, clear film.

Figure 2

First stimulus set -- used in Experiments 1, 2, 3 and 7.



Stimulus composition followed that of Taylor (1973) and Wales (pers. comm., 1974). The above array illustrates the Framed set (cf. Taylor).

The array of eight faces was fixed in the same order throughout, for both sets and for all subjects, following Taylor's design. (This order was originally randomly determined by Taylor.) Each face in the array was assigned a number (written on the back of the card) for ease of reference. The Figures illustrating the stimulus set show the order of the array and the reference-number of each face.

An additional set of five stimuli, together with five copies to serve as standards, were used for preliminary practice. These consisted of "stick figures" in various positions (standing, running, kneeling, lying down, kicking), drawn in black ink on square white cards, without frames.

Subjects

Subjects were children in attendance at a single playgroup in St. Andrews. The playgroup served both a local housing estate and professional families, so that the children came from a fair mix of backgrounds. Eight of the enrolled children were unavailable or unwilling to complete the experiment. This left a total of 18 subjects, assigned to two groups of nine, balanced as far as possible with respect to sex and chronological age. Group I contained five boys and four girls, whose ages ranged from 3;7 to 5;0 years (mean 4;2); Group II contained four boys and five girls, aged 3;7 to 4;9 years (mean 4;1).

Procedure

Within the playgroup room, a corner containing a low bench and table was partitioned off. Children were taken individually, invited to "come and look at some pictures". Child and Experimenter (E) sat side by side on the bench, and stimuli were laid on the table in front of the child.

The matching-from-sample task followed Taylor's method. One set of eight faces was displayed on the table. On each trial, a standard -- identical to one in the set -- was placed in the blank space at the centre bottom of the array. The child was instructed first to "look carefully at all of the pictures", then was asked, "Can you find a picture up here [E gestured to the array] that's just the same as this one [E pointed to standard]? Point to one that's just the same as this one" (with stress on "just the same"). If the child did not volunteer more than one choice as match-to-standard, E asked, "Is there another one that's just the same as this one?". The child's first and subsequent choices were recorded in order.

A test session continued for eight such trials, each of the eight stimuli of the array appearing once as standard. The eight standards were presented in random order, differing for each subject and each session. The child was given no feedback as to (in)correctness of his responses, although E offered frequent verbal encouragement throughout.

Each subject received two test sessions, one with Framed stimuli and one with Unframed. The nine children in Group I performed the task first with the Framed set, and about a week later with the Unframed; for Group II, the order was reversed. Thus each child acted as his own control across the two stimulus conditions.

A single test session lasted for about 10 minutes. It was preceded by five trials with the set of practice figures, using the test-session procedure. The practice array contained just two rows: three stimuli in the top row and two below, with a space in the middle of the lower row for the standard. If a child failed to choose a correct match on all of the first three practice trials, E interrupted with questions about the figures (What do you see in the picture? What is this man doing? Etc.).

E demonstrated and explained the correct choices for the first three standards, then presented the remaining two. Three children (not the youngest) failed in the first three practice trials, but all were correct on the last two of these trials.

Predictions

The children's responses yield three kinds of information:

(1) which pairs of faces are matched (correctly or wrongly) with each other;

(2) how many of the three attributes are actually correctly matched in those pairs;

(3) which of the attributes (Shape, Eyes, Mouth) is/are correctly matched.

For present purposes, the first two of these are only of passing interest. It is the last that is important, for the main question here concerns the effect of the stimulus condition as it relates to the contradictory results of Taylor and Wales (cf. Taylor's "ordering hypothesis" -- see earlier, p. 32). The expectation is thus as follows:
With the Framed faces (as Taylor's), there should be a greater number of correct matches of outer Shape of head (S) than of internal features Eyes (E) or Mouth (M). With the Unframed stimulus set (as Wales's), internal features should be correctly matched more often than Shape.

Now, each chosen face may actually match the standard in one of eight ways: in all three attributes, SEM (perfect match); in two attributes, SM, SE or EM; in just one attribute, S, M or E; or in Zero attributes (total mismatch). Thus responses can be categorised according to these eight sorts of matches (which, incidentally, inform about both the number and the kind of attribute(s) correctly matched).

The main test is then for effect of stimulus condition on the distribution of choices among these response categories. On the basis of the assumptions about what underlies the Taylor/Wales discrepancies, the predictions now become:

The presence of the square frame around the faces should create a bias in responding towards the SM, SE and S categories -- i.e., those in which Shape is correctly matched. The Unframed stimuli should be associated with responses mainly in the EM, M and E categories -- but also some in SM and SE, since these include an internal feature and some Shape-matching may occur incidentally, by chance. Both stimulus sets are expected to be equal with respect to the categories SEM and Zero -- total matches and mismatches.

In line with Taylor's data but of less importance here, we may also expect:

(a) that the greater the degree of actual similarity between particular stimuli, the more often they will be chosen as matching each other; e.g., that faces sharing two attributes with the standard will be chosen more often than those sharing only one attribute with it; and

(b) that younger children will achieve fewer perfect matches, and will choose on the basis of fewer actually-shared attributes, than older children.

RESULTS

Which (combinations of) attributes are correctly matched?

The numbers of choices falling into the eight response categories described above was calculated separately for the two subject groups and the two stimulus conditions. Table 1 summarises these results, for the children's first responses only. These data were arrived at as follows.

TABLE 1 (Experiment 1)

Distribution (per cent) of First Choices according to attribute(s) shared with the standard

ATTRIBUTES MATCHED		GROUP I (n = 9)		GROUP II (n = 9)		ALL SUBJECTS (N = 18)	
		Frame	UnFr.	Frame	UnFr.	Frame	UnFr.
2 attributes correct	SM	38.7	54.8	33.3	39.6	35.5	46.3
	SE	25.8	14.3	15.6	26.4	19.7	21.1
	EM	16.1	14.3	17.8	9.4	17.1	11.6
1 attribute correct	S	9.7	7.1	22.2	11.3	17.1	9.5
	M	9.7	7.1	8.9	7.5	9.2	7.4
	E	0	2.3	2.2	5.7	1.3	4.2
No. responses in above		31	42	45	53	76	95
No. Zero matches		1	0	4	0	5	0
No. SEM matches		40	30	23	19	63	49
Total responses		72	72	72	72	144	144

The total number of first responses per subject-group with each stimulus set is 72, i.e., nine subjects x eight trials each; the total number per stimulus set, combining both groups of subjects, is thus 144. In terms of the predictions set out above, the principal concern here is those response categories in which the choices match the standard in either two or one attribute(s) -- i.e., in which they shared with the standard SM or SE or EM; or S or M or E. So the numbers of perfect matches (SEM) and mismatches (Zero) were first subtracted from the total of 72 (or 144) responses. The remainder then gives the number of responses distributed among the six two- and one-attribute(s)-correct categories. In Table 1, the distribution over these six categories is expressed as a percentage of this remainder. By this means, any apparent differences due to differential numbers of Zero and SEM responses are eliminated, making the results comparable across the various subject groups and stimulus sets.

Now, look first at the right-hand columns in Table 1, where the results for the Framed and Unframed stimulus sets are compared with each other over all 18 subjects combined. (For the moment, we disregard the categories Zero and SEM.) We predicted that with the Framed set, the children's responses would fall mostly in the categories SM, SE and S. By inspection alone, this has indeed happened. For the Unframed set, in contrast, although some responses were expected in categories SM and SE, there should be more in EM, M and E both relative to the other categories within the Unframed-set condition and relative to the Framed set. Clearly, the responses do not follow the predicted pattern for the Unframed set. Although there is a slight shift in the expected direction for the one-attribute-shared categories S, M and E, for the two-attributes-shared ones the shift is in the opposite direction! That is, a greater number of choices match the standard in either SM or SE than in

EM; and the proportion of EM matches relative to SM or SE matches is lower for the Unframed than for the Framed set. The pattern is not essentially altered if the two groups of subjects are examined separately (left-hand columns of Table 1).

Nor is it altered by the inclusion in the data of the children's subsequent choices, as Table 2 shows. This table is constructed in the same manner as before, except for the obvious difference in total numbers of responses. While there were some refusals to make multiple responses, on most trials one or two additional choices were offered (with or without prompting), though occasionally there were as many as five further choices. (The bottom row of Table 2 adds the mean number of choices made per trial. The maximum possible in any one trial is of course eight, the number in the stimulus array.) This means that the total numbers of responses is no longer fixed at 72 per subgroup (or 144 per stimulus set with combined subject groups), but can vary. In practice, however, there was little difference in frequency of responding across the subject groups and stimulus sets (see second-bottom row in Table 2). (In both tables, of course, the maximum possible number of entirely correct matches -- category SEM -- remains at 72 per subgroup and 144 for combined groups.)

At first glance, then, the children's performance with Framed stimuli confirms Taylor's finding of Shape-based matching. But the Unframed set also appears to have elicited Shape-based matching, in contrast with Wales's claim of internal feature matching and Shape confusions with unframed stimuli.

In the form in which they have so far been presented, however, the data may be misleading for they contain several potentially complicating

TABLE 2 (Experiment 1)

Distribution (per cent) of All Choices (including multiple) according to attribute(s) shared with standard

ATTRIBUTES MATCHED		GROUP I (n = 9)		GROUP II (n = 9)		ALL SUBJECTS (N = 18)	
		Frame	UnFr.	Frame	UnFr.	Frame	UnFr.
2 attributes correct	SM	44.6	47.9	32.7	34.8	38.0	40.8
	SE	16.3	15.6	14.2	18.2	15.1	17.1
	EM	21.7	15.6	20.14	13.0	21.0	14.2
1 attribute correct	S	5.4	10.4	18.6	13.0	12.7	11.8
	M	6.5	9.4	9.7	11.3	8.3	10.4
	E	5.4	1.0	4.4	9.6	4.9	5.7
No. responses in above		92	96	113	115	205	211
No. Zero matches		5	4	11	7	16	11
No. SEM matches		59	57	44	37	103	94
Total responses		156	157	168	159	324	316
Mean no. choices to each standard		2.2	2.2	2.3	2.2	2.3	2.2

factors, apparent in the above tables. Firstly, it is evident that the assumption concerning the equality across conditions of numbers of perfect matches and mismatches was unfounded: both more SEM and more Zero responses occurred with Framed than with Unframed stimuli. Secondly, regardless of stimulus set, the subjects in Group II fared generally worse than those in Group I, as reflected in the differential achievement of SEM matches. These two factors together mean that the distributions of choices among the "crucial" categories are based on unequal response rates across the various subgroups and conditions, which may help to conceal other real differences in response patterns. Thirdly, the partitioning of the categories itself complicates the issue. Recall that one of the predictions, following Taylor, was that there should in general be more choices of two-attribute matches than of one-attribute matches. This is confounded with the primary expectation of SM, SE and S matches with Framed stimuli and of EM, M and E matches with Unframed: rather than choosing single M or E matches, the children may have been selecting matches of SM and SE instead.

Apart from these considerations, statistical analyses of the above data are unwarranted because of the small or empty cell frequencies. It was therefore decided that it may be more profitable to collapse the response categories, into simply correct matches of Shape, of Mouth and of Eyes. Accordingly, the total number of times each attribute was correctly matched was summed across trials and children for each subgroup, and converted into a percentage of the total possible correct. Table 3 presents these data for first choices only. (Since each child makes eight first choices per stimulus set, the total number of times an attribute can possibly be correctly matched in a subgroup of nine subjects is 72 -- or 144 for both groups combined.)

TABLE 3 (Experiment 1)

Per cent correct matches of each attribute in First Choices

ATTRIBUTE MATCHED	GROUP I* (n = 9)		GROUP II* (n = 9)		ALL SUBJECTS** (N = 18)	
	Frame	UnFr.	Frame	UnFr.	Frame	UnFr.
SHAPE	87.5	86.1	76.4	83.3	81.9	84.7
MOUTH	83.3	86.1	69.4	68.1	76.4	77.1
EYES	73.6	59.7	54.2	56.9	63.9	58.3
Total no. possible correct per attribute	72	72	72	72	144	144

*Percentage calculated out of 72 = Total no. first choices = Total no. possible correct per attribute

**Percentage calculated out of 144 = Total no. first choices = Total no. possible correct per attribute

TABLE 4 (Experiment 1)

Per cent* correct matches of each attribute in All Choices (including multiple)

ATTRIBUTE MATCHED	GROUP I (n = 9)		GROUP II (n = 9)		ALL SUBJECTS (N = 18)	
	Frame	UnFr.	Frame	UnFr.	Frame	UnFr.
SHAPE	80.8	81.5	70.8	71.1	75.6	76.3
MOUTH	76.9	80.9	65.5	66.0	71.0	73.4
EYES	59.6	56.1	48.8	52.8	54.0	54.4
Total no. responses	156	157	168	159	324	316
Total no. possible correct per attribute	288	288	288	288	576	576

*Percentage calculated out of Total no. responses made in each condition separately ($\neq$ Total possible correct).

Table 4 gives the results for all choices (first and subsequent choices combined). Again, they are expressed as percentages, but this time not out of the total theoretically possible correct matches, for this far exceeds the total possible in practice. For each subgroup, the maximum possible correct per attribute could be taken as 288 (i.e., for any given standard, there exist four possible correct matches to each separate attribute; 9 subjects x 8 trials x 4 possible correct = 288). But in fact, to obtain such a maximum, a subject would have to make seven choices to each standard (since the array of eight contained only one entire mismatch for each standard)! Typically, the number of responses to a single standard was actually only two or three (see the bottom two rows of Table 2). The percentages quoted in Table 4 were therefore calculated on the basis of these actual response frequencies, which were treated as "equivalent in practice" to the maximum possible correct.

Once more, inspection alone suggests that while performance with the Framed stimuli exhibited the expected trend, performance with the Unframed ones did not. In fact, the two stimulus conditions produced remarkably similar response patterns.

These data were subjected to analyses of variance ("Three-factor analysis with repeated measures, Case I": Winer, 1971, pp. 539ff). Each subject was observed under all combinations of Factors B and C (two stimulus sets and three attribute response categories, respectively), but under only a single level of Factor A (order of presentation of stimulus sets -- Framed first or Unframed first). The subject-groups were nested under Factor A but crossed with Factors B and C. Summaries of the results of the analyses are presented in Table 5 (for first responses only) and Table 6 (for all responses).

TABLE 5 (Experiment 1)

Summary of analysis of variance on First Choices (see text for explanation)

SOURCE	SS	df	MS	F	p
Between Subj.	91.05	17			
A (order)	22.23	1	22.23	5.17 (1,16)	<0.05
error a	68.81	16	4.30		
Within Subj.	266.83	90			
B (stim. sets)	0.08	1	0.08	<1	
AB	2.08	1	2.08	1.35 (1,16)	>0.25
error b	24.67	16	1.54		
C (attributes)	60.02	2	30.01	7.16 (2,32)	<0.005
AC	2.35	2	1.18	<1	
error c	134.96	32	4.22		
BC	2.17	2	1.08	<1	
ABC	3.17	2	1.58	<1	
error bc	37.33	32	1.17		

TABLE 6 (Experiment 1)

Summary of analysis of variance on All Choices

SOURCE	SS	df	MS	F	p
Between Subj.	768.49	17			
A (order)	44.08	1	44.08	<1	
error a	724.41	16	45.28		
Within Subj.	1294.50	90			
B (stim. sets)	0.23	1	0.23	<1	
AB	1.56	1	1.56	<1	
error b	278.70	16	17.42		
C (attributes)	306.69	2	153.34	8.758 (2,32)	<0.005
AC	4.06	2	2.03	<1	
error c	560.26	32	17.51		
BC	0.57	2	0.29	<1	
ABC	5.57	2	2.79	<1	
error bc	136.85	32	4.28		

Whether all responses or just the first ones are considered, these analyses confirm that there was virtually no effect whatsoever of stimulus set ($F < 1$ in both cases): the children behaved in the same way with the Unframed as with the Framed faces. As might be anticipated, the effect for responses is highly significant ($p < 0.005$ in both cases): the frequency of correct matches differed significantly among the three attributes.

For a more precise examination of this effect, *t*-tests for matched pairs were conducted on the differences between pairs of attributes in frequencies (per cent) of correct matches. The results follow:*

Shape vs Mouth: There were no differences in correct matches between Shape and Mouth. This was true for both stimulus sets, and for first and multiple responses ($p > 0.1$ minimum).

Shape vs Eyes: Throughout, Shape was correctly matched significantly more often than Eyes. Framed set, first choices: $t = 3.56$, $p < 0.005$; all choices: $t = 4.26$, $p < 0.001$. Unframed set, first choices: $t = -4.78$, $p < 0.001$; all choices: $t = -4.39$, $p < 0.001$. (Note that the negative values of *t* for the Unframed set show that the outcome was the opposite of that predicted: there *should* have been more Eyes- than Shape-matches, according to the original expectation.)

Eyes vs Mouth: These tests gave less consistent results. For Framed stimuli, the expectation from Taylor's findings was for more Eyes- than Mouth-matches. While no such difference emerged for first responses ($t = -1.82$, $p > 0.1$), there was a significant difference when all responses were considered ($t = -3.76$, $p < 0.001$) -- and in the "wrong"

---

* Unless otherwise indicated, the p-values are for one-tailed levels, since the expected direction of difference was specified in advance; df = 17 throughout.

direction: Mouth was correctly matched more often than Eyes. There was no indication from Wales as to direction here. Two-tailed tests revealed significant differences for the Unframed stimuli in both first responses ($t = 2.21$, $p < 0.05$) and all responses ($t = 3.21$, $p < 0.01$). Again, Mouth was matched correctly more often than Eyes.

Returning to the analyses of variance, Tables 5 and 6 show that none of the interaction effects reaches anything near significance. One factor remains, however: the effect of the order of presentation of the two stimulus sets is significant at $p < 0.05$ for first responses -- though it disappears when subsequent responses are included. But this effect is probably not as troublesome as it might seem, since order of presentation does not interact with the other factors. From Table 3, it is clear that the response patterns themselves -- the crucial issue here -- varied little under the two orders of presentation. Rather, the matching performance of Group I subjects (who received the Framed set first) was, in overall terms, superior to that of Group II (who received the Unframed set first). It is not a practice effect, for Group I subjects did generally worse in their second session (with Unframed stimuli) than in their first. This is demonstrated in Tables 7 and 8 (first responses and all responses, respectively), showing the number of choices that were correct in all three, in two, in one, or in no attributes, regardless of type of attribute.* Possibly there is something about the

---

* The visual search patterns of the children are possibly relevant here. During the task, the Experimenter made informal notes on how individual subjects went about scanning the array. After looking first at the standard, some children proceeded around the array, moving from one face to another without further recourse to the standard. Others looked from standard to one comparison face, and back to the standard again before looking at another comparison face. The informal observations suggest that it was those children who made frequent reference to the standard who were more likely to choose a perfect match, or at least a two-attribute match. Moreover, children who did not refer to the standard while searching for a first choice as match to it often did not even

TABLE 7 (Experiment 1)

Number of attributes correctly matched in First Choices

NO. OF ATTRIBUTES CORRECT	GROUP I (n = 9)		GROUP II (n = 9)		ALL SUBJECTS (N = 18)	
	Frame	UnFr.	Frame	UnFr.	Frame	UnFr.
(SEM) THREE	40	30	23	19	63	49
(SM/SE/EM) TWO	25	35	30	40	55	75
(S/M/E) ONE	6	7	15	13	21	20
(Zero) NONE	1	0	4	0	5	0
Total responses	72	72	72	72	144	144

TABLE 8 (Experiment 1)

Number of attributes correctly matched in All Choices

NO. OF ATTRIBUTES CORRECT	GROUP I		GROUP II		ALL SUBJECTS	
	Frame	UnFr.	Frame	UnFr.	Frame	UnFr.
(SEM) THREE	59	57	44	37	103	94
(SM/SE/EM) TWO	70	76	72	76	142	152
(S/M/E) ONE	22	20	41	39	63	59
(Zero) NONE	5	4	11	7	16	11
Total responses	156	157	168	159	324	316

Unframed stimuli that provoked more confusion errors (and perhaps carried over for Group II subjects to influence their later performance with the Framed stimulus set).*

Some brief comments on Taylor's other hypotheses are in order here. Tables 7 and 8 confirm that, of those choices that were not perfect matches, a greater number shared two attributes with the standard than shared just one attribute with it (and there were comparatively few complete mismatches). It appears that the children did indeed operate on the basis of <u>degree</u> of similarity as well as <u>kind</u> (in Taylor's words, <u>order</u>) of similarity.

Also in keeping with Taylor's results, the younger children in the present sample attained fewer correct matches than the older. This is apparent from Table 9 (for first responses only; the picture is the same when multiple responses are included); the entire sample of 18 subjects, cutting across the original groups, was simply divided into the nine younger (ages 3;7 - 4;2, mean 3;11 years) and the nine older (ages 4;4 - 5;0, mean 4;8).**

/continued/ return to it when asked for further choices to the same standard; their gaze moved on around the array, restarting from the preceding choice face. Thus in multiple responses to a single standard, such a child was likely to choose a face (horizontally, vertically or diagonally) <u>adjacent to his preceding choice</u>. The relationship between matching performance and visual search patterns in tasks requiring the scanning of an array clearly deserves further investigation.

* This explanation is preferred to the alternative that Group II and Group I subjects genuinely differed in ability, since Unframed stimuli again generated more errors than did Framed in Experiment 3 (see pp. 85, 96-7). But some independent measures of the children's IQ would have been helpful here.

** There was little difference between ages in overall frequency of multiple responding. Mean choices per standard ranged from 2.1 to 2.3 across the subgroups of age x stimulus set. This response rate includes both extra choices made on the Experimenter's prompting and extra choices offered spontaneously -- the former outnumbered the latter. The rate is considerably lower than that reported by Taylor. His two youngest subject groups (of similar ages to the present sample) offered 5.0 and 3.5 choices per standard -- and that spontaneously alone.

TABLE 9 (Experiment 1)

Number of attributes correctly matched by Younger and Older subjects (First Choices only; Subject Groups I and II combined)*

NO. OF ATTRIBUTES CORRECT	FRAMED		UNFRAMED	
	YOUNGER	OLDER	YOUNGER	OLDER
THREE	24	39	17	32
TWO	31	24	44	31
ONE	14	7	11	9
ZERO	3	2	0	0
Total responses	72	72	72	72

*Both of the original subject groups are combined here. The Younger subjects are 5 from the original Group I and 4 from Group II; the Older are 4 from Group I and 5 from Group II.

TABLE 10 (Experiment 1)

Number of correct matches of each separate attribute by Younger and Older subjects (First Choices only; Subject Groups I and II combined)

ATTRIBUTE MATCHED	FRAMED		UNFRAMED	
	YOUNGER	OLDER	YOUNGER	OLDER
SHAPE	54	64	63	59
MOUTH	52	58	50	61
EYES	42	50	37	47
Total matches	148	172	150	167

Despite the difference between age groups in numbers of attributes correctly matched, age does not appear to have affected the response patterns in terms of kind of attribute(s) correctly matched -- see Table 10 (again for first responses only, since subsequent choices followed the same pattern).

With regard to his "ordering hypothesis" (see p. 32), Taylor found a considerable degree of consistency within individuals, and was able to classify his subjects according to order of similarity (= confusion errors) among the attributes. In the present experiment, response patterns of individual subjects proved not particularly illuminating. They are presented and compared with Taylor's findings in Appendix 2.I.

One final point needs to be made about the children's responses in Experiment 1. Frequency of correct matching differed not only among the three attributes, but also among the eight faces themselves. Table 11 shows the number of times each face was correctly chosen as match to its standard in first responses. (The eight faces are numbered 1 to 8; which number applies to which face is given on p. 37, duplicated in Appendix 1, Figure 1. The total number of times any face can be chosen correctly is of course 18, the number of subjects.)

Faces 5, 6, 7 and 8 were correctly matched the most often with both stimulus sets, accounting between them for twice as many correct matches as faces 1 to 4. The difference in correct matches between these two groups of faces, 1 to 4 versus 5 to 8, reached significance when t-tests for matched pairs were applied (for Framed stimuli, $t = 5.531$, $p < 0.001$; for Unframed, $t = 2.496$, $p < 0.05$; two-tailed values with df = 17 in both cases). But Table 11 should be viewed alongside Table 12, which shows the number of times each face was chosen whether it was a correct

TABLE 11 (Experiment 1)

Number of times each face in the array was correctly matched to its standard (First Choices only)

	STIMULUS FACE NUMBER								RANK OF FACES BY "CORRECTNESS"
	1	2	3	4	5	6	7	8	
No. correct matches when Framed	4	4	6	4	10	8	14	13	7, 8, 5, 6
No. correct matches when Unframed	6	3	4	5	9	7	7	8	5, 8, 6 & 7
Total correct matches per face	10	7	10	9	19	15	21	21	7 & 8, 5, 6
	total for Faces 1-4 = 36				total for Faces 5-8 = 76				

TABLE 12 (Experiment 1)

Number of times each face in the array was chosen, correctly or incorrectly (First Choices only)

	STIMULUS FACE NUMBER								RANK OF FACES BY FREQUENCY OF CHOICE
	1	2	3	4	5	6	7	8	
No. choices* when Framed	7	11	17	11	21	23	33	21	7, 6, 5 & 8
No. choices* when Unframed	14	9	19	16	22	20	21	23	8, 5, 7, 6
Total no. choices per Face	21	20	36	27	43	43	54	44	
	total for Faces 1-4 = 104				total for Faces 5-8 = 184				

*Out of a total of 144 theoretically-possible occasions per cell match or not. Clearly, it is these same faces, 5 to 8, that were chosen the most often -- and again, significantly more than faces 1 to 4 (for Framed stimuli, $\underline{t} = 3.878$, $p < 0.002$; for Unframed, $\underline{t} = 2.122$, $p < 0.05$; two-tailed with $df = 17$). Presumably, then, the larger proportion of correct matches for faces 5, 6, 7 and 8 is pretty well accounted for by their higher frequency of selection to <u>any</u> standard. Since it does not bear directly upon the main issues of the present investigation, further discussion of this finding and the results of a supplementary test for preferences among the eight faces are removed to Appendix 2, Section II.

<u>Summary and conclusions</u>

The results of Experiment 1 broadly support Taylor's general conclusions. The children did make confusion errors, but by no means randomly. For one thing, about 80% of their choices were either perfect matches or two-attribute matches (87% for older subjects with Framed stimuli). So these preschool children must have had a reasonably well developed understanding of sameness, for they were clearly basing their decisions to a considerable extent on <u>degree</u> of physical similarity.

It is equally clear that their decisions rested on kind as well as on number of shared attributes. Overall, Shape was correctly matched in about 80% of first responses, Mouth in about 75%, and Eyes in only about 60%. That is, the confusion errors were consistently (and significantly) biassed towards Eyes. As Taylor notes, such a systematic tendency to overlook one particular attribute argues against the view that young children simply have an "immature" understanding of what "just the same" means. Rather, their confusion errors stem either from an inability to discern a particular feature of difference, or from a deliberate disregard for it.

The present results confirm Taylor's also in that the attribute most often correctly matched was Shape. They disagree, though, over the internal attributes: for Taylor's subjects, Mouth ranked poorest, whereas in the present case Mouth was correct almost as often as Shape. This discrepancy is especially apparent in the comparison of results for individual subjects (as opposed to pooled responses), given in Table A2-1 of Appendix 2. For those of the present subjects who could be classified with respect to which attribute predominated in matches, the most common order was the one in which Mouth was correctly matched above all else; so the numbers of Shape matches must have been made up by unclassified subjects, who showed inconsistent or minimal-error patterns.

The results with the Unframed stimuli, however, contradict Wales's claim. Where internal features were expected to be correctly matched more often than Shape, the reverse actually obtained. Both Framed and Unframed faces elicited remarkably similar patterns of confusion errors, in fact.

The starting hypothesis was thus not substantiated. The assumption that the differential matching-patterns observed between Taylor and Wales stemmed from Wales's omission of Taylor's square frame around the stimuli proves to have been unfounded. What then could account for their contradictory findings?

There was in fact a second major, methodological difference between Taylor's and Wales's procedures. The present Experiment 1 followed Taylor's matching-from-sample task. Wales, in contrast, used the pair-comparison method: only two stimuli were presented at a time, and the children had to judge whether the two were the same or not. There are several reasons for supposing that the two methods might elicit different sorts of behaviour (although it is not clear _a priori_ what specific differences in responses might be expected):

(1) In matching-from-sample, the entire stimulus set is on view all the time. In pair-comparison, only two stimuli are present on any trial. Both Taylor and Wales (1970; also Taylor, 1973) and others (e.g., Garner, 1966; Tversky 1977) have emphasised the importance for sameness judgments of the subject's awareness of the physical structure of the stimulus set, and in particular of the structural relations among stimuli. If the child is to use such structural-relational information, in the pair-comparison situation he has to hold in memory previously-presented pairs and build up an internal cognitive construction of the stimulus set piece by piece over the course of several trials -- obviously a complex task for a young child. In matching-from-sample, he has all the necessary information in front of him and need not rely on memory nor on the internal construction process. (Cf. Over & Over's [1967] finding of fewer confusion errors in sameness judgments under matching-from-sample than under pair-comparison type conditions.)

(2) Matching-from-sample requires a _multiple_-choice decision; the decision is a far simpler, two-way one in pair-comparison. This carries with it at least two implications for task performance:

(a) Visual scanning requirements. We saw earlier that young children may use immature visual scanning strategies when inspecting stimulus figures. As Taylor (1973) notes, it is possible that in matching-from-sample the child may fail to search the entire array before making his choice. (For studies of the order in which children inspect members of an array of stimuli, see Gottschalk _et al_., 1964; Santostefano & Paley, 1964; Elkind & Weiss, 1967; see also pp. 53 & 55, footnote.)

Thus he may stop at the first stimulus encountered that is something like the standard. This type of visual search behaviour does not operate in pair-comparison.

(b) Response requirements. The kind of response required in matching-from-sample could be called an "active" one: searching the array for the correct match and then pointing to it. Not only is the search component absent in pair-comparison, but the required response is a more "passive" one: merely answering "yes" (they are the same) or "no" (they are not), without active stimulus selection. This introduces a potential sub-problem: as Taylor comments, young children are reputed to exhibit a bias towards answering "yes" in such contexts, whereas the pair-comparison format in this instance requires the answer "no" considerably more often than "yes" for correct performance.

If it is indeed the case that the Taylor/Wales discrepancies can be explained on methodological grounds, then which of the differences between the two types of task is the critical one?

That the two types of task do indeed yield different results is supported to some extent by Taylor's work. After the matching-from-sample task, he gave some of the original subjects a pair-comparison task with the same set of schematic faces (1973, Experiment IV). To summarise his findings:

(1) Numbers of attributes confused. The children now confused stimuli that actually differed in two attributes as often as, if not more often than, those differing in just one attribute, and a relatively high proportion of complete mismatches were judged to be the same. (This could perhaps be accounted for by a bias towards "yes" answers.) The finding agrees with that of Over and Over (1967): pair-comparison seems to be a harder task for young children than matching-from-sample.

(2) Kind of attribute(s) confused. Taylor's regular pattern of Shape-correct and Eyes-confused, observed under matching-from-sample, had disappeared (see his Table XXIII). Now, for pairs judged to be the same but which actually shared only one attribute, the errors were about equally distributed among the attributes. For pairs actually sharing two attributes, the tendency was for the greatest numbers of confusions to occur over Eyes, with Mouth next, and Shape more likely than the other features to be correctly matched. Compared with the matching-from-sample response patterns, while Shape was again the attribute most often correctly judged, the order was reversed for Eyes and Mouth. In the pair-comparison data, however, error proportions were relatively high over all attributes, and in general there was no really clear order among them.

In any case, these data still do not correspond with Wales's finding of internal feature matching and Shape confusions. In terms of the next research step, two possibilities were considered. One was to attempt pair-comparison with both Framed and Unframed stimuli. The second was to try comparison not within pairs but within triads of stimuli. In the latter method, usually known as the three-choice-oddity problem, three unidentical stimuli are presented on a trial, from which the child is asked to choose two that resemble each other (e.g., "Which two go best together?").* While the child's responses in such a task tell little about his ability to judge sameness in the sense of perfect identity (since no triad contains an identical pair), they should reveal the basis on which his judgments are made, i.e., which attributes he matches and confuses.

---

* For the sake of brevity, this method will here be called simply the triads task.

Triadic presentation resembles pair-comparison in that only a part of the stimulus set is available at any one time. But unlike pair-comparison, the required response involves some degree of visual search among stimulu and "active", multiple-choice selection. The method of triads, then, can help to tease out which of the differences between pair-comparison and matching-from-sample might be important. The question is whether the response patterns that emerge with the triads method will correspond with those of the matching-from-sample task, or those of the pair-comparison.

EXPERIMENT 2

Triads

Materials

The stimuli were the cards that were used as "standards" in Experiment 1: two sets each of eight schematic faces, one set Framed and one cut out and Unframed.

In his pair-comparison task, Wales presented the stimuli not on a flat surface but held up by hand. Pilot tests with 12 children revealed no obvious differences in responses whether the Unframed faces were held up, or presented flat on the matt grey background of Experiment 1, or on white or multi-coloured backing sheets. The matt grey background was therefore retained for the Unframed set in Experiment 2. In each triad with the Framed stimuli, the three square cards were laid in a horizontal row, nearer sides touching. With the Unframed set, the three faces were separated by the same distance as in the Framed triads, but were displayed against the grey background and with no surrounding frames.

A total of 56 triads, each of three unidentical stimuli, is generated from the eight stimuli per set. Each subject performed the task in two sessions of 28 trials each (both with the same stimulus set), separated by one to two weeks. To determine the order of presentation, each triad was randomly assigned a number from 1 to 56. Two response sheets were compiled for each subject, one listing triads 1 to 28, the other triads 29 to 56. Half the subjects received triads 1 - 28 in their first test session; the other half received triads 29 - 56 first. Within each response sheet, the order of the trials was randomised, a different order used for each subject. Over the course of the 56 triads, each of the eight faces appeared 21 times -- seven times in each location within the triad, left, centre or right. With this proviso, the positions of the stimuli in each triad were determined randomly, but were constant over all subjects.

For practice, the set of five "stick figures" used in Experiment 1 was retained, but now supplemented by four identical isosceles triangles outlined in black ink on unframed square cards. Two of these cards were white and two pink. (These additional cards were supplied by Wales; he and Taylor had used them as stimuli in previous research.)

Subjects

From an initial sample of 44 children, 36 were retained as subjects.* Their backgrounds were various. Permanently resident in St. Andrews were

---

* From the original sample, eight subjects were dropped: four because they did not complete the task; two because they responded apparently at random and gave idiosyncratic reasons for their responses; and one because it was discovered that English was not his first language. One final subject, chosen at random, was dropped from the Unframed Group in order to balance the numbers in the two subject groups.

nine in attendance at a playgroup, 18 in the first class of a school, and five children of lecturers in psychology at the university. A further four came from Glasgow families on holiday in St. Andrews, who volunteered in response to an advertisement. The subjects were assigned to two groups, counterbalancing age and sex as far as possible. Both groups contained ten boys and eight girls. The Framed Group, age range 3;5 to 7;0 years (mean 5;5), received only the Framed set of stimuli; the Unframed Group, ages 3;4 to 7;2 (mean 5;5), received only the Unframed set.

Procedure

As before, all children were taken individually, invited to come and look at some pictures. In the playgroup, a corner of the hall was sectioned off for the purpose. In the school, subjects were taken to the vacant staff room. Lecturers' children were seen either in their own homes or in the University Psychological Laboratory, and the four Glasgow children were brought to the Experimenter's home. With the exception of three of the children tested in their own homes, during the task the child and the Experimenter sat side by side on chairs and stimuli were placed on a table in front of the child. For the remaining three, child and Experimenter sat together on the floor and the stimuli were laid on the floor.

On each trial, the three appropriate stimuli were laid out, and the child was asked, "Which two look most the same?".* After making his

---

* Although the wording of this question sounds strange to adult ears, the children appeared to have no difficulty with it. Pilot work had determined that other forms (e.g., "Which two go together best?") elicited responses based not on shared visual features but on, e.g., position or some idiosyncratic or seemingly random basis. This was evident from the children's reasons for their choices: "Because this one's near", "Because he just went for a walk", etc.

choice, he was asked, "How are they the same?" (and occasionally, "Can you see anything about them that's <u>not</u> the same?"). The Experimenter avoided using the word "face". The pair chosen from the triad was recorded, along with the child's comments in reply to the further questions about their sameness/difference. (Those comments will be discussed in Part 4.)

The first test session was preceded by at least five practice trials, using three triads from the set of stick figures and two with the triangles. On the first trial with the stick figures, two of the stimuli were identical and the third different. Subsequent practice triads contained no identical pairs. A stick-figure triad might then be grouped as, say, two upright (standing, running) and one lying down, while the triangles varied in colour of card and orientation of triangle. As in test trials, the child was asked for the two most the same and how they were the same. Additionally, to ensure that he understood the task and was aware of possible alternative bases of choice, he was asked (a) what each of the three pictures showed, (b) how the chosen two differed from each other, and (c) how the other two possible pairs in the triad were the same and different. The children performed well in practice, although some -- especially the younger ones -- were initially hesitant to give verbal descriptions and reasons. In these cases, the Experimenter prompted until the child answered himself, and extra practice trials were sometimes given, up to a total of seven. Before the second test session, the child received just four practice trials, two each with the stick figures and triangles.

During the test sessions, no feedback was given, only frequent encouragement. A single session lasted for about 20 minutes.

Expectations

The experiment was of the "let's see what happens" type; no specific directional hypotheses were proposed. The primary concern was which of the three attributes -- Shape, Eyes, Mouth or configurations of them -- would be most often shared (matched) in the face pairs chosen from the triads. More specifically, the question was which of three main directions the triads results would follow:

(a) Response patterns for both stimulus sets resembling that of the matching-from-sample task: There would be more choices in the categories denoting Shape matches than in those denoting Eyes or Mouth matches.

(b) Patterns for both stimulus sets resembling that of Wales's pair-comparison task: There would be more choices involving matches of internal features than of Shape.

(c) Patterns differing between the two stimulus sets: Framed stimuli would produce mainly Shape matches (cf. Taylor/matching-from-sample), while Unframed stimuli would produce internal feature matches (cf. Wales/pair-comparison).

There is, of course, another possibility: that the responses would be about equally distributed among the three attributes, i.e., that the triads results would parallel those of Taylor's pair-comparison task.

RESULTS

The children's responses were first examined in terms of the categories that were used to classify the matching-from-sample choices in Experiment 1: SM, SE, EM, S, M, E and Zero, denoting which (and how many) attributes were shared in the chosen pair. Since no triad contained any actually identical pair, the former category SEM, standing for a perfect match in all three attributes, was no longer applicable.

Table 13 summarises the results, comparing the response distributions for Framed and Unframed stimuli. In each category, the mean number of choices per subject was calculated, for the two subject groups separately (n = 18 in each). The maximum number of choices per subject that could fall into any one category was 24, out of the total of 56 face pairs selected by the child. The left side of Table 13 presents these means. On the right, the same data are expressed as percentages of all choices, for ease of comparison with the results of the matching-from-sample task given in Tables 1 and 2 (pp. 42 and 45).

It is clear that the patterns of responses in the triads task incline towards the third one of the possible directions they were expected to take. That is, they move in the direction originally predicted, but not actually found, for the matching-from-sample data (see pp. 40-41). To recap: in the Framed condition, most responses were expected to fall in Categories SM, SE and S; in the Unframed, there should be a shift away from these Shape match categories towards the internal feature ones, EM, M and E. Now, in the triads task, there is a striking cross-over effect (marked in Table 13) over the Framed and Unframed conditions between the categories SE (Shape + Eyes the same) and EM (Eyes + Mouth same); and between the single-attribute-shared categories Shape and Mouth. It appears that where the subjects of the Framed Group chose primarily Shape matches with Mouth second, the Unframed Group chose Mouth matches first and Shape next. Both groups, then, chose matches of Shape + Mouth (SM) about equally often (on at least 70% of the 24 trials per subject in which a SM match was possible), and single matches of Eyes equally infrequently.

The cross-over between Shape and Mouth is confirmed if the responses are pooled across categories to give simply the total numbers of matches

TABLE 13 (Experiment 2)

Distribution of responses according to attribute(s) matching in the chosen face-pairs

ATTRIBUTES SHARED		MEANS PER SUBJECT (max. poss. per cell = 24)		PER CENT OF ALL CHOICES (based on a total of 1008 responses)	
		FRAMED	UNFRAMED	FRAMED	UNFRAMED
two attributes matched	SM	16.8	18.5	30.1	33.0
	SE	13.8	7.2	24.7	12.8
	EM	7.9	13.8	14.1	24.7
one attribute matched	S	9.0	4.9	16.1	8.8
	M	4.7	9.6	8.3	17.1
	E	3.0	1.6	5.4	2.9
	Zero	0.8	0.4	1.4	0.7

TABLE 14   (Experiment 2)

Per cent matches of each attribute overall

ATTRIBUTE SHARED	FRAMED	UNFRAMED
SHAPE	70.8	54.7
MOUTH	52.5	74.8
EYES	44.2	40.4

of Shape, of Mouth and of Eyes. Table 14 presents these data. As with the corresponding matching-from-sample results in Tables 3 and 4, the new triads data are expressed as percentages of the total number of times an attribute could possibly be matched. Now, a child may match a given attribute 56 times (i.e., on every trial); for each attribute, the maximum number of matches possible is thus 1008 within each group of 18 subjects.

The percentages in the triads data appear lower than those for the matching-from-sample. This is because the responses are no longer independent from each other across the attribute-categories. In matching-from-sample, it is possible for a child (one who achieves perfect matches throughout) to obtain 100% for each of Shape, Mouth and Eyes matches. With the triads format, perfect matches are not possible; a child can match 100% of the time on one attribute, but if he does so, the maximum he can achieve on the other two attributes is 66.7% each. Because of this interdependence among the attributes, the kind of analysis of variance performed on the matching-from-sample data of Experiment 1 could not be applied here. Instead, a series of t-tests was conducted -- though it should be recognised that these are still not independent from each other, in that a significant value in one test necessarily entails a converse difference elsewhere. Since the expected directions of any differences were left open at the outset, all tests were two-tailed.

Little of interest emerged from comparisons between pairs of attributes (by t-tests for matched pairs) within the Framed and Unframed conditions separately. The likely reason is that most of the choice pairs (about 70%) shared two attributes. In neither condition was there a difference between Shape and Mouth in numbers of matches. With the Framed stimuli, Eyes and Shape matches differed significantly (t = 2.820, df 17, $p < 0.02$); similarly, the Unframed stimuli produced a significant difference between Eyes and Mouth (t = 4.304, df 17, $p < 0.002$). This outcome simply restates the finding that Eyes matches occurred relatively infrequently with either stimulus set, while Shape and Mouth were both matched fairly often throughout.

Comparison of the matches of each attribute between the Framed and Unframed conditions revealed one significant outcome: a difference in the frequency of Mouth matches (t = 2.758, df 34, $p < 0.01$). Thus the cross-over effect appearing in Tables 13 and 14 is to some extent confirmed. This lends support to the claim that (at least with the method of triads) omission of the frame around the faces creates a shift towards an internal feature as the basis of matching.

In contrast with the matching-from-sample results, there was almost no difference between younger and older children in the number of attributes matched per choice in the triads task. As noted earlier, two-attribute matches were chosen about two-thirds of the time or more, by all subjects.* The reason is probably that most children, regardless of age, clearly made their choices on the basis of one particular attribute, consistently matching that attribute above the other two throughout the task. For example, suppose a child strictly chooses pairs that share Mouth. Then his responses will be about equally distributed among the three categories SM, EM and M; i.e., he must choose twice as many two-attribute matches as one-attribute matches even if he is not aware that Shape or Eyes is shared as well as Mouth. Given the tendencies, noted earlier, (a) generally to choose two-attribute matches over single-

---

* However, all but one of the complete mismatch pairs were chosen by the younger half of the sample. But since there was a total of only 21 mismatches in all (1% of the responses), this finding is of negligible size.

attribute matches where possible, and (b) to choose the SM configuration specifically, then the proportion of single-attribute matches may be further reduced. So a strict Mouth-matcher's total of 56 responses could be distributed as 24 SM matches, 16 EM matches and 16 M matches (Mouth being matched on every trial, and Eyes or Shape only incidentally).

In fact, 13 of the 36 subjects produced just such a completely strict pattern. Relaxing the strictness criterion to allow a 10% deviation from the "dominant" attribute (i.e., to include subjects who matched by their dominant attribute on at least 50 trials) revealed a further 11 clear matching patterns. Table 15 presents the numbers of subjects who produced clear patterns, by stimulus condition and dominant attribute matched.* From this table, the Mouth- and Eyes-matchers were then collapsed into a single subgroup to yield a 2 x 2 contingency table -- Framed versus Unframed, and matchers of Shape versus matchers of internal features -- on which a Fisher test was performed. The null hypothesis can be rejected at the 0.05 level (two-tailed). Thus the two stimulus conditions produced significantly different matching patterns: most classifiable subjects in the Framed Group chose by Shape, while most in the Unframed Group chose by an internal feature, Mouth.

Notice that the children's matching patterns were not affected by the division of the task into two separate sessions. The mixed pattern of the inconsistent subjects was not due to their switching from one

---

* Six children in each stimulus condition were inconsistent. None came near the 50-trial criterion; the most achieved on any one attribute was 44 matches. There seems to be a rather nice dichotomy here: children who matched the same attribute throughout were <u>clearly</u> consistent, while children who showed a mixed pattern were <u>clearly</u> inconsistent. Note that this rate of 33% subjects "unclassifiable" agrees with that observed in Experiment 1 (see Appendix 2.I).

TABLE 15 (Experiment 2)

Distribution of subjects among classes according to their "dominant attribute" matching patterns

CLASS OF SUBJECT	FRAMED GROUP (n = 18)	UNFRAMED GROUP (n = 18)
SHAPE-MATCHERS	9	3
MOUTH-MATCHERS	3	8
EYES-MATCHERS	0	1
MIXED-MATCHERS	6	6

attribute to another between the two sessions. Only one child clearly switched from one attribute to another; he did so half-way through the first session, and continued matching the second attribute in his later session.

In comparison with the matching-from-sample results, the clarity of the response distributions found here -- particularly when individual subjects' matching patterns are examined -- suggests that the triads design may be especially suitable when the main concern is the kind of attributes matched, rather than the number of attributes. On the basis of this experiment, systematic response patterns seem more likely to emerge with the triads -- and when they do, they are readily spotted. Given a subject who consistently matches one particular attribute on most trials, it can also easily be seen which (if either) of the other two attributes he matches secondarily, for the following reason. Consider a strict Mouth-matcher again. On eight trials out of the 56, all three faces share the same level of Mouth. (This happens equally, on different sets of eight trials, for the other two attributes.) If the child consistently chooses the EM configuration on these trials, we have determined that he matches Eyes second to Mouth and overlooks Shape. In practice, what usually happened on such trials was that the child initially pointed to all three faces as the same. The Experimenter then asked, "Can you find two that look more the same than the others?"; a pair was usually readily selected in response. This hints that, in at least some cases, the non-dominant attributes were being deliberately ignored but could be brought into play when "necessary".

Another situation arose in which the Experimenter had to press for choices of a single pair. Two of the older subjects (one aged 6;4 years, from the Framed Group, and one of 7;2 in the Unframed) had difficulty in deciding which two faces to choose. Often they made a first choice on the basis of a particular attribute, but then reconsidered it. When this happened, they proceeded to give extremely detailed descriptions of the ways in which all three possible pairs in the triad were the same and different. If more than one pair shared two attributes, they resisted making a final choice (sometimes becoming quite upset) until the Experimenter emphasised that they should choose the two that looked more the same than the others. (The final choice was the one taken for the response data in such instances.) Over the course of the task, these two children switched between choosing on the basis of one particular attribute and choosing on the basis of number of attributes shared, thus making a relatively high proportion of two-attribute matches. Six adults given this triads task with the Framed stimuli responded on the basis of number of shared attributes, and also refused when this system yielded more than one pair in a triad as possible answers (e.g., when more than one pair had two attributes in common). When prompted, they would then choose by kind of attribute (the favourites among the adults again being Shape and/or Mouth shared). It seems, then, as if the two children were in transition between the childlike pattern of matching a particular dominant attribute and the adultlike comparison of numbers of shared attributes. The impression from these children's comments was that they became upset because they could see both ways of responding and did not know which was the "correct" one in terms of the Experimenter's expectations; only constant encouragement would induce them to proceed. Given the ages of these children, it is tempting to speculate on a possible association between their behaviour in this task and the supposed onset of the Piagetian stage of concrete operations. This could be an interesting avenue for further research.

Summary and conclusions

In the matching-from-sample task of Experiment 1, the children behaved the same way whether the stimuli had a surrounding frame or not: in both conditions, they matched primarily on the basis of Shape, with Mouth next. Now, in the triads task, a difference has begun to emerge. Framed stimuli were again chosen mostly by Shape, with Mouth next; but Unframed stimuli were matched more according to form of Mouth, with Shape next. We are brought full circle back to the initial hypothesis that the presence of a surrounding frame would lead to matches based on the outer Shape of the faces, while in the absence of the frame the subjects' matching behaviour would be determined by the internal features Mouth and/or Eyes. But the original notion of what underlay the discrepancy in the Taylor and Wales findings is shown to have been too simplistic. It is not just that matches of the Unframed faces were chosen on the basis of an internal feature -- this only happened in the context of triadic comparison.

In the introduction to Experiment 2 (pp. 61ff), some of the differences among the three methods, matching-from-sample, pair-comparison and triads, were discussed in terms of the nature of the stimulus array and the sorts of behaviours they call for. Matching-from-sample and triads tasks, recall, are alike in that an active selection response is required, based on a judgment among multiple alternatives. They differ in that the entire stimulus set is on display throughout the matching-from-sample task, while only a subset is available at any one time in the triads. This suggests that it may be the stimulus array, and concomitantly the nature or extent of the visual search behaviour involved, that is important in determining whether or not the Unframed stimuli elicit matching of internal features.

Since Wales actually employed a pair-comparison rather than a triads format, the next step in the attempt to account for the differential response patterns observed by Taylor and Wales was obvious: to present the Framed and Unframed faces for pair-comparison.

EXPERIMENT 3

Pair-comparison

Materials

The eight Framed and eight Unframed stimuli of the triads task were used again. Two further matching sets were constructed, since identical pairs were now required (unlike in the triads task). Stimuli were presented side by side. For those with frames, the inner edges of the two square cards touched each other. Unframed pairs were displayed the same distance apart as the Framed faces, but were laid on the grey backing sheet as previously. Additional practice stimuli were the stick figures and triangles used in the triads task.

Subjects

Sixty-four subjects were tested in all. The two youngest (ages 3;2 and 3;5 years) were subsequently rejected, one because he answered that the two faces were not the same on <u>every</u> trial even when they were identical, and one because he answered "same" throughout.* This left 62

---

* Presumably, these two youngest subjects failed to understand the task or the instructions. They were the only seemingly age-related failures in the entire series of experiments. Perhaps other kinds of tasks are needed to probe very young children's notions of sameness -- although subjects of just over three years old managed pair-comparison in Experiment 8.

subjects, drawn from a variety of sources: 13 from a playgroup in St. Andrews and 10 from one in Oxford; 13 from the first class of a primary school in St. Andrews and 26 from one in Leuchars. They were assigned to two groups of 31, counterbalancing age, sex and source as far as possible. The Framed Group contained 18 boys and 13 girls, age range 4;0 to 6;11 years (mean 5;3); they received only the Framed stimuli. The Unframed Group, receiving only Unframed stimuli, contained 16 boys and 15 girls, ages 3;9 to 6;11 years (mean 5;2).

Procedure

Subjects were again invited one at a time to come and look at some pictures. In the playgroups, a table and chairs were set up in a corner; in the schools, a separate room was made available and similarly equipped. As usual, child and Experimenter sat side by side and stimulus pairs were laid on the table in front of the child.

A single set of eight faces generated 36 possible pairs: eight identical matches and 28 unidentical. A different random order of presentation of the 36 trials was given to each subject. Within pairs, stimulus positions were also randomised over subjects, with the proviso that each face appear on the left and on the right equally often over the course of the task.

On each trial, a pair was presented and the child was asked, "Are these two (pictures) the same or not the same?". "Not the same" was used rather than "different" in view of the possibility that young children's sense of the meaning of "different" may be poorly developed, especially in relation to their understanding of "same" (e.g., Donaldson & Wales, 1970; Webb et al., 1974). The Experimenter also avoided using the word "face". On some trials (arbitrarily selected and differing across subjects), the child was also asked, "How are they the same?" and/or "How are they not the same?". In the triads task, it had been attempted to question the children in this way on every trial. Since it considerably lengthened the session and caused attention-wandering in the crucial decision part of the trials, however, such questioning was now limited. Some children nevertheless offered the information spontaneously on other trials. The subject's decisions and comments about sameness-difference were recorded.

Five preliminary practice trials proved sufficient for all children except the youngest two who were later dropped from the sample. Two of these trials used the stick figures -- one identical and one unidentical pair. Three used the triangles: One identical pair, one differing in colour of card, and one differing in both colour and orientation of the triangle. Throughout practice, the child was pressed for descriptions of the stimuli and reasons for his decision, until the Experimenter was satisfied that he understood the task and was aware of the various relevant stimulus features.

The task was performed in a single session, lasting for about 15 minutes. Frequent encouragement was given, but no feedback as to (in)correctness of responses.

Predictions

The eight identical-pair trials are of little interest here save as a check on the child's general understanding of the task and instructions. There is no reason to expect incorrect judgments on these trials unless very occasionally (e.g., through temporary attention-lapse), and correct judgments tell nothing about the main present concern -- which attribute(s) the child correctly and wrongly matches. The focus will thus be on the 28 unidentical pairs.

In view of the triads results, the pair-comparison method, which likewise presents only a part of the stimulus set at a time, may be expected again to produce matches of an internal feature with the Unframed stimuli -- as Wales reported. From Experiments 1 and 2, there is also no a priori reason to suppose that the Framed set should not continue to elicit decisions based primarily on Shape. But Taylor's finding in his pair-comparison task (1973, Experiment IV) of no clear pattern among the attributes casts some doubt after all on what to expect, at least with the Framed stimuli. Perhaps the general supposed tendency of young children in such situations to answer "yes" (they are the same) even in the face of the large proportion of "no" responses required for correct performance could counteract any effect of attributes.

While the possibility was recognised that confusions might turn out to be about equally distributed over the three attributes (as Taylor found), directional predictions were nonetheless made: for correct judgments of Mouth and/or Eyes and confusions of Shape by the Unframed Group; and conversely, for correct judgments of Shape and confusions of internal features by the Framed Group.

RESULTS

First, the "control" trials in which both stimuli were identical. Only one subject in the Framed Group and three in the Unframed made errors here -- one each. This total is so minute that these trials can be safely ignored.

Secondly, six subjects in each group made correct judgments on every trial. In the Framed Group, the youngest of these was aged 4;2 years, then one of 5;10 and the others over six years. Correct subjects in the Unframed group were one of 4;10 years, then four spread over age five, and one of 6;10. It is possible, then, for some children to perform perfectly well on this task at a quite early age, and yet for others of nearly seven years to make errors. Since it is errors that produce the response patterns of interest here, the 12 entirely correct subjects are excluded from the following discussion unless otherwise noted. This leaves 25 error-producing subjects in each condition.

As before, the main concern is which attribute(s) the child does and does not detect/use as the basis of sameness. The findings in the first two experiments were couched largely in terms of correct matches. With the pair-comparison format, however, the comparable information is obtained by looking not at correct judgments but at <u>incorrect</u> ones. This is because of the nature of the response required for correct performance on all 28 critical trials: the answer "not the same". Suppose, for example, that a child makes correct ("not the same") judgments of all pairs actually sharing SE, E or S, and also makes wrong ("same") judgments of all pairs actually sharing just SM, EM or M. From his <u>error</u> distribution (SM, EM, M) we can see at a glance that <u>Mouth</u> is what he notices or uses, since it is the only attribute common to all his error categories, and that any concomitant Shape or Eyes matches are merely incidental -- he is, in the terminology developed in the discussion of Experiment 2, a "strict Mouth-matcher". This is not so immediately obvious from his <u>correct</u> judgments.

Throughout the discussion of the pair-comparison results, then, expressions like "the subjects judged by Mouth", "Mouth was correctly matched", etc., are synonymous with "the subjects made <u>errors</u> on trials in which the pair shared the <u>same Mouth</u>". For the sake of clarity, the terms "error" and "confusion" are here distinguished: <u>error</u> refers to <u>trials or face-pairs</u> eliciting a wrong "same" response; <u>confusion</u> refers to an actually differing attribute that is wrongly treated as the same. Thus a child who makes errors on Mouth-sharing pairs judges predominantly by Mouth and confuses Shape/Eyes.

Let us now look at the distribution of error-responses over the familiar attribute categories: SM, SE, EM, S, M, E and Zero. Table 16 presents this. A child can make a maximum of four errors in any one category. Since we are dealing with 25 subjects in each group, the maximum number of errors possible per cell is 100. The figures given in the left half of Table 16 can thus be read equally as simply the total numbers of errors actually made, or as the percentage of times an error was made when it was possible to do so, for each response category. On the right side of the table, the same data are expressed in terms of the proportion falling to each category of all errors actually made per subject group, for comparison with the results of the matching-from-sample task given in Tables 1 and 2 (pp. 42 and 45) and of the triads task in Table 13 (p. 70).

Table 17 shows the results when the categories are collapsed, as previously, into simply all correct judgments of Shape, of Mouth and of Eyes. The table is divided into two parts. On the left, the data from Table 16 have simply been recast, as follows. The basis is which attributes are actually shared in face pairs that were wrongly judged to be the same. In Table 16, a maximum of four such errors per subject was possible in each response category. The new collapsed categories each sum across three of the former ones: Shape = SM + SE + S; Mouth = SM + EM + M; Eyes = SE + EM + E. The maximum wrong judgments per subject now is thus 12 per category. For example, in Table 16 terms, a strict Shape-matcher's responses would appear as 4 SM, 4 SE and 4 S. Now, they would be distributed as 12 S, 4 M and 4 E. In each stimulus condition,

TABLE 16 (Experiment 3)

Distribution of responses (= errors) according to attribute(s) shared in face-pairs wrongly judged to be the same*

ATTRIBUTE(S) SHARED		TOTAL ERRORS PER CELL		PER CENT OF ALL ACTUAL ERRORS	
		FRAMED (n = 25)	UNFRAMED (n = 25)	FRAMED (n = 25)	UNFRAMED (n = 25)
two attributes correct	SM	62	64	35.6	28.7
	SE	37	17	21.3	7.6
	EM	36	76	20.7	34.1
one attribute correct	S	15	6	8.6	2.7
	M	15	51	8.6	22.9
	E	8	9	4.6	4.0
	Zero	1	0	0.6	0
		Total no. of errors made (max. poss. = 700):		Percentages above calculated on basis of error totals:	
		174	223	174	223

*Correct "not the same" judgments of non-identical face-pairs are not of interest here; hence the data in this table represent errors. The table shows what attributes were actually shared (correctly matched) when wrong "same" responses were given, in order to show what formed the basis of sameness judgments. By inversion, the confused attributes are arrived at; e.g., the high proportion of SM errors (Shape & Mouth shared) indicates a high proportion of Eyes confusions.

TABLE 17 (Experiment 3)

Per cent correct judgments of each attribute when errors were possible

ATTRIBUTE CORRECTLY JUDGED	For correctly-matched attributes in pairs receiving wrong "same" judgments		For correctly-discriminated attributes in pairs receiving correct "not the same" judgments	
	FRAMED GROUP (n = 25)	UNFRAMED GROUP (n = 25)	FRAMED GROUP (n = 31)*	UNFRAMED GROUP (n = 31)*
SHAPE	38.0	29.0	87.9	72.6
MOUTH	37.7	63.7	87.7	93.6
EYES	27.0	34.0	81.3	75.6
percentages based on a total possible for each cell of:	300	300	496	496

*The two columns on the right-hand side include the six subjects per Group who were correct on every trial. These entirely correct subjects have been omitted until now.

the maximum possible number of entries per attribute is thus 300 (25 subjects x 12 entries). The left side of Table 17 expresses the data as the percentage actual entries per attribute out of the 300 possible.

These figures, however, are not readily comparable with the corresponding matching-from-sample data presented in Tables 3 and 4 (pp. 47-48) nor the triads data in Table 14 (p. 71). To facilitate this cross-reference, the right side of Table 17 "inverts" the data. Since the results from the first two experiments used correct matches, not errors, the basis now is which attributes actually differ in face pairs that were correctly judged to be not the same. The strict Shape-matcher, for instance, is correct on all pairs differing in S, SE, SM and SEM (corresponding to the former shared-attributes-categories EM, M, E and Zero, respectively). Thus each new attribute category now sums across four of the former ones, and the maximum possible entries for each attribute is 16 per subject. Now, also, the 12 subjects who were entirely correct need to be included, bringing the maximum possible per attribute within a stimulus condition to 496 (31 subjects x 16 categories). This formed the denominator for the percentages shown on the right side of Table 17.

Whichever way the data are represented, the outcome is much the same. The Framed Group made correct judgments of Shape and Mouth equally often, and of Eyes less often. The Unframed Group clearly judged predominantly by Mouth and confused Shape and Eyes both.

Statistical tests confirmed these apparent patterns. Analysis of variance could be performed since, unlike in the triads data, entries in one attribute category did not affect those in another -- a child could in theory contribute maximally to all categories independently. The method employed was the two-factor analysis with repeated measures on one factor (Winer, 1971, pp. 514ff). Each subject was observed under all levels of Factor B (three attribute categories, Shape, Mouth, Eyes) but under only one level of Factor A (two stimulus sets, Framed, Unframed). Subjects were thus crossed with Factor B but nested under Factor A. The results are summarised in Table 18. While there was no broad difference in responses between the Framed and Unframed Groups, the effect of attributes was marked ($F = 13.160$, df 2, 96, $p < 0.001$). The stimulus condition x attributes interaction effect was also highly significant ($F = 7.049$, df 2, 96, $p < 0.005$). That is, the two groups of subjects made roughly equal amounts of errors overall, but their responses were distributed differently among the three attributes.

The nature of these differences was further examined by series of _t_-tests, first on pairs of attributes within each stimulus condition (df = 24 for these matched-pairs tests). For the Framed set, there was no difference between Shape and Mouth responses, but Shape was matched significantly more often than Eyes ($t = 2.205$, $p < 0.025$; both tests were one-tailed, since it was predicted at the outset that Shape would predominate over both of the internal features with Framed stimuli). The test between Mouth and Eyes just failed to reach significance ($t = 2.037$, $p > 0.05$) at the _two_-tailed level (no directional difference was predicted). For the Unframed set, judgments were based on Mouth far more often than on Shape, as predicted ($t = 5.793$, $p < 0.001$, one-tailed); but Eyes and Shape responses did not differ (one-tailed). There was also a marked difference between Mouth and Eyes ($t = 4.101$, $p < 0.002$, _two_-tailed level since no direction was specified).

Secondly, _t_-tests were performed on each attribute category between the two stimulus conditions. Since directions of expected differences were specified at the outset, all p-values here are for one-tailed levels;

TABLE 18 (Experiment 3)

Summary of analysis of variance on distribution of responses among the attributes Shape, Mouth, Eyes

SOURCE	SS	df	MS	F	p
Between Subj.	619.04	49			
A (stimulus set)	34.56	1	34.56	2.833 (1,48)	>0.1
Subj. within groups	584.48	48	12.18		
Within Subj.	987.33	100			
B (attributes)	190.49	2	95.25	13.160 (2,96)	<0.001
AB interaction	102.04	2	51.02	7.049 (2,96)	<0.005
B × Subj. within groups	694.80	96	7.24		

df = 60. In Shape matches, the Framed Group far outdid the Unframed ($\underline{t}$ = 7.023, $p < 0.001$). For both Mouth and Eyes the differences just reached significance ($\underline{t}$ = 2.038 and 2.249 respectively, $p < 0.05$ for both), the Unframed Group making more matches here, as expected.

The patterns within each stimulus condition and the differences between the two are highlighted when we look at individual subjects instead of pooled responses, as Table 19 shows. The following criteria were established for classification of subjects.

(1) <u>One-attribute matchers</u>. Keeping to the terminology introduced for the triads results, <u>strict</u> matchers of one particular, dominant attribute were those who made a total of exactly 12 errors, i.e., made a false "same" judgment on all face pairs in which that attribute was shared, but on no other trial. Thus, as already noted, the pattern for a strict Shape-matcher was 4 SM, 4 SE, 4 S (summing to 12 S, 4 M, 4 E in the collapsed categories). In the Framed Group, two subjects produced such strict distributions, one matching Mouth, the other Shape. In the Unframed Group, there were five strict Mouth-matchers.

To include subjects whose patterns were obviously robust but not 100% strict, the criterion was relaxed. To be classed as a <u>clear</u> one-attribute matcher, the subject had to make at least 10 errors involving matches of the dominant attribute; also, one error was permitted on a face pair that did not share the dominant attribute. The top three rows of Table 19, excluding the figures added in parentheses, show the numbers of subjects who met these requirements. The "least strict" pattern admitted here was one of 4 SM, 4 SE, 2 S, 1 EM (= 10 S, 5 M, 5 E in the collapsed categories) -- still a clear Shape-matcher.

Additionally, however, there were several subjects who fell short of the above criteria but who nonetheless obviously matched one particular attribute over the others. Since their failure to attain the status of "clear matcher" resided in small number of errors rather than in mixed matching pattern, they were treated separately from those who were unclassified because of inconsistency. A subject was called a weak one-attribute matcher if: there were at least eight entries (out of the possible 12) over the three appropriate dominant-attribute-shared categories and no more than one outside them; and at least two entries appeared in each of the dominant attribute categories; and in total the dominant attribute was matched at least twice as often as either of the others. In Table 19, these subjects have been added in parentheses beside the clear one-attribute matchers. For all five weak matchers, Mouth was the dominant attribute (typical patterns being 2 SM, 3 EM, 3 M and 3 SM, 4 EM, 2 M, 1 Zero). Their bias towards Mouth was considered strong enough to justify this treatment.

(2) Two-attribute matchers. Several subjects matched by two attributes simultaneously; neither alone was sufficient, but the presence of matches of both together invariably elicited a wrong "same" judgment. A strict two-attribute matcher was therefore one who made exactly four errors, all falling in a single two-attribute category (i.e., scoring just 4 SM, or just 4 SE, or just 4 EM). For present purposes, a subject was not excluded if he made only one other error outside that category; but within the category, all four errors had to occur.

(3) Unclassified subjects were:

(a) those who made too few errors to allow a clear pattern to be seen. With the exception of the clear two-attribute matchers just described, any child making seven or fewer errors was treated as unclassifiable.

TABLE 19 (Experiment 3)

Distribution of subjects among classes according to their dominant attribute matching patterns (see text for explanation)

CLASS OF SUBJECT BY ATTRIBUTE(S) MATCHED		FRAMED GROUP (n = 31)	UNFRAMED GROUP (n = 31)
one-attribute matchers	SHAPE	3	1
	MOUTH	2 (+2)*	9 (+3)*
	EYES	2	2
two-attribute matchers	SHAPE & MOUTH	5	0
	SHAPE & EYES	2	0
	MOUTH & EYES	0	3
unclassifiable because no dominant attribute	INCONSISTENT	1	2
	TOO FEW ERRORS	8	5
	ALL CORRECT	6	6

*Additional numbers in brackets are the "weak Mouth-matchers"

(b) those who were inconsistent. Three subjects made 11 or more errors but not exhibiting any dominant-attribute pattern.

(c) those who were all correct. To complete the subject groups, the six children in each who made correct judgments on all 28 trials were added.

The top two rows isolated from Table 19 form a 2 x 2 contingency table (Shape- versus Mouth-matchers, Framed versus Unframed stimuli) on which a Fisher test was performed. The alternative to the null hypothesis was the original, directional prediction that the two stimulus conditions would elicit different response distributions, with the Framed Group matching predominantly Shape and the Unframed matching an internal feature. Whether or not the weak Mouth-matchers were omitted, no significant difference obtained at the one-tailed level. Nothing would be gained by incorporating Eyes-matchers since these would simply further reduce the difference between the two groups.

It is clear from the various statistical presentations that only part of the predictions made at the start of Experiment 3 were confirmed. As expected both from the present triads results and from Wales's earlier pair-comparison finding with Unframed stimuli, subjects given a pair-comparison task with Unframed faces made sameness judgments primarily on the basis of a shared internal feature (Mouth). Further, for the first time in the present series of experiments, Eyes were not the most frequently confused attribute -- the Unframed Group confused Shape most, thus strengthening their tendency towards internal feature matching.

The Framed Group did not behave as expected from the triads results; that is, Shape no longer predominated outright over the internal features. The Framed Group pattern tended, rather, to agree with Taylor's corresponding results. As noted in the introduction to Experiment 2, the clear distributional bias found by Taylor in matching-from-sample, of Shape most often correct and Mouth least, disappeared almost completely in his pair-comparison task. In the present pair-comparison task, either Mouth matching was as common as Shape matching or, indeed, for several subjects it was only the Shape-plus-Mouth combination that counted. In contrast with Taylor's finding, however, Eyes matches were still significantly less frequent.

But this apparent contradiction is likely explained by the fact that in pair-comparison Taylor's oldest subject was 5;7 years, while the present sample extended to 6;11. To examine this possibility, the present subjects were divided on a post-hoc basis into the 16 younger and 15 older in each condition, the upper limit of the younger subgroups being 5;2 years. From Table 20, it is evident that the younger subjects in the Framed condition made about equal numbers of matches of all attributes, i.e., behaved just as did their peers in Taylor's experiment. It was the older children who exhibited a dominance of Shape and/or Mouth. For the Unframed Group, no such difference in the response distributions of the two age groups appeared, only a general difference in error frequency.

With regard to this last point, however, there was a marked departure from Taylor both in the overall frequency of errors observed and in the numbers of attributes confused. Taylor's overall error frequencies have been added to Table 20; his subjects produced a far higher error rate than did the present ones.* In fact, both his groups tended to make

---

* There is a possible explanation for the difference in overall error frequencies between the two experiments. Where subjects in the present pair-comparison task were asked, "Are these two (pictures) the same or not the same?", Taylor asked his, "Is this the same as this?" (pointing to the stimuli). In the former question, the negative is explicitly stated

TABLE 20 (Experiment 3)

Distribution of responses according to attribute(s) shared in face-pairs wrongly judged to be the same, by Younger and Older subjects (cf. Taylor, 1973, Table XXIII).

ATTRIBUTE(S) SHARED		FRAMED Younger (n = 14)*	FRAMED Older (n = 11)*	UNFRAMED Younger (n = 15)*	UNFRAMED Older (n = 10)*
two attributes correct	SM	27	35	45	19
	SE	28	9	12	5
	EM	21	15	44	32
one attribute correct	S	8	7	6	0
	M	7	8	38	13
	E	8	0	4	5
	Zero	0	1	0	0
Total no. errors		99	75	149	74
% error rate		22.1%	17.9%	33.3%	17.6%
Taylor's % error rate **		Group V 57.1%	Group IV 30.0%		

* These n's exclude the 12 subjects who were all correct. In total, there were 16 subjects in each Younger Group and 15 in each Older Group.

** Calculated on the basis of information presented in Taylor (1973). His Group V had 10 subjects, aged 3;10 - 4;8, mean 4;2 years; his Group IV had 13 subjects, aged 4;3 - 5;7, mean 4;10. In the present sample, the ages were:
Younger: Framed -- 4;0 - 5;2, mean 4;7. Unframed -- 3;9 - 5;2, mean 4;6.
Older: Framed -- 5;3 - 6-11, mean 6;0. Unframed -- 5;2 - 6;11, mean 6;0.

more errors on pairs sharing only one attribute than on pairs sharing two. Almost 50% of all his errors fell to two-attribute differences, as against almost 40% to pairs differing by one attribute. The correspon-ding figures for the present pair-comparison results with the Framed stimuli were 23% for two-attribute differences and 77% for one-attribute differences (with no effect of age). Further, whereas Taylor's two subject groups made surprisingly high proportions of errors (12 and 14%) on pairs actually differing in all three attributes, only one wrong "same" response to a Zero pair was produced in the whole of the present experiment. As in the matching-from-sample and triads tasks, then, the present subjects clearly took into account degree as well as kind of actual similarity. (Had they been using purely the kind of attribute shared, their errors would have been distributed as 33% on two-attribute differences and 67% on one-attribute differences.)

Remaining for a moment with the proportion of errors to correct judgments, notice that overall the Unframed stimuli elicited slightly fewer correct judgments than the Framed (although not significantly so, according to the analysis of variance summarised in Table 18). This accords with the observation in the matching-from-sample task of Experi-ment 1 that the Unframed Set received fewer complete SEM matches than the Framed. The agreement lends some support to the earlier suggestion that there may be something about the Unframed faces that makes fully

---

/continued/ as an alternative to the "same" judgment. Indeed, the fear had been that the children might reply "not the same", showing a recency-of-mention effect; but pilot tests revealed no such tendency. Taylor's question form, on the other hand, could well lead the children -- who were young enough to be not yet fully conversant with the niceties of language -- to guess that the answer "yes" (they are the same) was expected. While this could have operated to produce Taylor's generally high error rate, however, it is unclear how it could explain his finding of a greater number of errors on pairs differing by two attributes than on pairs dif-fering by only one attribute.

correct matching performance less likely with them. To put it another way, the child given Unframed stimuli seems more likely to adhere to a single dominant attribute and consistently overlook the others.

Summary and conclusions

The series of experiments reported here began with the assumption, based on extensive evidence from other sources, that young children make confusion errors when they choose or compare visual stimuli for sameness. That confusions would occur was taken for granted here. The question was what kinds occur -- whether regular patterns can be discerned of systematic correct matching of certain stimulus features and confusions of others; and if so, what perceptual/cognitive factors might explain them.

Specifically, the present programme was stimulated by prior contradictory observations as to which attribute(s) children took as the basis of sameness among schematic faces that could vary in Shape, Eyes and Mouth. Taylor (1973) found that Shape was the most often correctly matched attribute, whereas Wales reported (pers. comm., 1974) confusions of Shape and better performance with the internal features. The starting objective, then, was to attempt to account for these differential results, as a means to shedding more light on the nature and process of children's sameness judgments and confusion errors.

One specific perceptual factor immediately presented itself as a candidate. The most salient difference between the two experiments seemed, initially, to lie in the fact that in Taylor's stimulus set each face was surrounded by a square frame, while in Wales's the frames were missing and instead each face was cut out around its elliptical head-shape. On a number of grounds, it was reasonable to suppose that such an external perceptual feature should affect which part of the stimulus was

was noticed or used. That the presence of a surrounding frame would elicit correct judgments of the outer Shape of the faces more than of the internal face features, while its absence would lead instead to correct performance on internal features rather than on Shape, thus seemed the obvious first prediction to test. It was falsely supposed, however, that presence/absence of a frame was the sole major difference between Taylor's and Wales's procedures. Subsequently, it emerged that, whereas Taylor had employed the framed faces in a matching-from-sample task, Wales presented his unframed set for pair-comparison.

Over the course of the first three experiments of the present study, then, both the perceptual and methodological factors were examined for their relative contributions to frequency and distribution of correct decisions and confusions among the three stimulus attributes. Throughout, two identical sets of eight schematic faces were used, one Framed (like Taylor's) and one Unframed (like Wales's). Three types of task were administered: matching-from-sample (like Taylor's), triads (novel) and pair-comparison (used by both Wales and Taylor). They were selected to vary amount of the total stimulus set available at a time, type of visual search behaviour involved, type of decision and response required, in an attempt to untangle which particular aspects of the methodological differences might be important. Table 21 indicates where the commonalities and differences in these task variables lie among the three methods. Under each method, the findings are summarised for the two stimulus conditions separately. The corresponding findings of Taylor (Framed only) and Wales (Unframed only) have been added in parentheses where appropriate.

An interaction between the task and stimulus variables is evident. With the Framed faces, Shape predominated as the basis of sameness in matching-from-sample and triads, but not in pair-comparison. This

TABLE 21

Summary of experimental manipulations and corresponding results for Experiments 1, 2 and 3.

TASK VARIABLES		TYPE OF TASK: MATCHING-FROM-SAMPLE	TRIADS	PAIR-COMPARISON
Amount of stimulus set on view	all	✓		
	part		✓	✓
Visual search/ Decision basis	scan array & compare multiple pairs	✓	✓	
	compare single pair			✓
Response type	active stimulus selection	✓	✓	
	passive verbal yes/no answer			✓
STIMULUS VARIABLES		RESULTS (Attributes ranked according to correctness of judgments)*		
Framed (cf. Taylor)		$S \geqslant M > E$ $(S > E > M)$	$S > M > E$	$S = M > E$ $(S = M = E)$
Unframed (cf. Wales)		$S \geqslant M > E$	$M > S > E$	$M > E = S$ $(M ? E > S)$

* Entries in parentheses were the patterns observed by Taylor or Wales

Key: ">" -- "is correctly judged more often than"; "=" -- "is correctly judged as often as"; "?" -- order-relationship between the attributes is unknown suggests that active, multiple-choice stimulus selection is a necessary task component for the frame to produce a bias towards Shape matches. When the response required merely passive judgments of just two alternatives, the distribution of correct matches among the three attributes levelled out. But whether all or just part of the stimulus set was on display at a time seems to have been irrelevant.

Against this, for the Unframed stimuli to produce decisions based on matches of an internal feature, the crucial requirement is apparently that only a few stimuli be on view. Provision of the entire stimulus array in fact led to a near reversal, from Mouth-based to Shape-based responses. And in contrast with the Framed condition, the type of response involved now made only a slight difference.

Notice, however, that independently of the task x stimulus set interaction, the matching-from-sample format elicited a Shape bias regardless of presence/absence of frame; and also that the pair-comparison format, relative to the other methods, does seem to have generally shifted the basis of responding further towards internal features with both stimulus sets. It is possible that the type of array in which the stimuli were displayed in the matching-from-sample task made Shape generally more salient, overriding any differential effect of the Framed/Unframed condition; and conversely, that the pair-comparison display tended to lead to a focus on internal features, counteracting the influence of the frame.

With regard to his two sets of results, Taylor remarked on the perfect achievement of correct "same" judgments on actually identical pairs in pair-comparison, contrasting with matching-from-sample in which identical matches could be missed. He offered two explanations for this pair-comparison effect. One concerned the young child's possible bias towards "yes" answers, mentioned earlier. While this could also account for his own high error rate, the present finding of a rather low proportion of false-positive responses argues against any major effect of such a bias in the pair-comparison replication (see also footnote on pp. 94 and 96).

In addition to the "yes" bias, Taylor suggested that presentation of only two stimuli at a time highlights their relevant aspects, minimising effects of poor visual scanning and distraction that could happen with the large array. Again, such a possibility could explain the correct "same" judgments. But, as Taylor himself comments, it is unclear how either factor could be involved in the errors, which he found to occur slightly more frequently on pairs differing by <u>two</u> attributes than on single-difference pairs. Neither suggestion, without qualification, accounts successfully for the different response distributions obtained with Framed and Unframed stimuli in the present experiment, particularly in view of the consistent bias towards Mouth in the Unframed condition. While both a "yes" bias and the notion that pair-comparison display highlights the important aspects of the stimuli are intuitively plausible explanations, in practice they do not seem to play much of a role. We are left simply with the conclusion that, for some as yet unknown reason(s), the pair-comparison method seems to create a general shift towards matching by an internal attribute. The triads task, in contrast, allowed the Framed Shape-dominance and Unframed Mouth-dominance patterns to emerge at their most strongly differentiated levels.

There is little direct evidence from other sources on the effects of methodology. The test requires holding stimulus set constant and varying the type of task. This has, however, been done with the Gibson sets of letter-like forms and their transformations, which have been used in several studies.

In the original work (Gibson, Gibson, Pick & Osser, 1962), the task was a kind of visual matching-from-sample in which the standards and an array of transformations were on view throughout (see earlier, pp. 20-21, for a sketch of this work). To recap on the main relevant results: topological transformations (such as gap-to-close) were the least likely to be confused, line-to-curve and rotation or reversal were next, and perspective transformations were frequently confused. The same pattern was found when the stimuli were real letters of the alphabet.

Subsequently, Anne Pick (1965) used some of the Gibson stimuli with kindergarten-age children, in three types of task. The first copied the original visual matching-from-sample method and yielded much the same error pattern as before. The second and third tasks used the tactual mode in a pair-comparison format. In the second task, the two stimuli were received successively, while in the third they were received simultaneously. Now, line-to-curve transformations elicited more confusions than did rotation/reversal. But it is not clear whether it was the type of stimulus presentation (pairs instead of a large array) or the difference in mode of stimulus perception (tactual instead of visual) that produced this result. The latter seems probable -- i.e., that line-to-curve variations became harder to distinguish by touch.

Pick was also testing for the children's use of a criterial attribute or, instead, their use of a prototype in these tasks. The matching-from-sample results were taken to indicate that it was subjects given the opportunity to use criterial attributes who made the fewest errors, although there was some suggestion that a certain amount of prototype learning also occurred. In the tactual pair-comparison task with _successive_ stimulus presentation, there was no difference between the criterial attribute and prototype conditions. But with _simultaneous_ pair-comparison, subjects given the chance to use criterial attributes were clearly superior, and the prototype group now made as many errors as control subjects whose pretraining had emphasised neither criterial attributes nor prototypes. Thus it is possible that the memory component imposed by the second of Pick's three tasks affected the child's detection and/or use of stimulus features. (Cf. Over and Over's [1967] finding that diagonal lines were more often confused under successive than under simultaneous pair-comparison conditions, and that matching-from-sample gave the fewest errors.)

In all of the Gibson and Pick work, however, subjects were _trained_ on the task. What was being measured was in essence transfer to new standards or to new transformation types. Schaller and Harris (1974) again used the Gibson letter-like forms, but now in visual pair-comparison and without prior training. Despite the difference in task and training conditions, their response patterns showed a striking similarity to those of Gibson _et al_. (1962). One point of divergence, however, was in Schaller and Harris's finding of an _increase_ (instead of a decrease) in rotation/reversal errors with age -- an observation that accords with some other experimenters' findings with other types of stimuli (e.g., Williams _et al_., 1976). Schaller and Harris ascribe Gibson _et al_.'s observed reduction of errors here to the original researchers' emphasis on pre-training on reversals.

Outside the domain of letter-like forms, there is some evidence of a task effect from Hale and Lipps (1974). Stimuli were a square, circle or triangle, coloured orange, yellow or blue. Two types of task were administered. One was a triads matching task similar to the present one of Experiment 2. Hale and Lipps pointed out that the triads method forces a choice between the stimulus attributes, so that one attribute tends clearly to predominate in the children's responses. To determine to what extent the children actually detected the non-dominant attribute as well, they next presented the same stimuli to the same subjects in a component-selection task. First, the child learned the spatial locations of stimuli in a triad. In this phase, the two stimulus attributes, shape and colour, were redundant with each other, e.g., the square was always yellow or the circle always blue. Those three stimuli were then turned face away from the child and the test phase, involving a sort of matching-from-sample, began. The child received a stimulus composed of only one attribute -- just a coloured card, or a plain card showing a geometric form. He had to point to the location in the original triad of the one matching the test card. He thus had the chance to show he could match either stimulus attribute equally well.

Subjects were divided into two groups by age: roughly 3½ to 4½ years, and 4½ to 6½ years. In the first triads task, about two-thirds of the younger group matched by colour, while almost all of the older subjects matched instead by shape. The component-selection task showed the same pattern as did the triads for the younger children, although there was some evidence that the non-dominant attribute of the triads was now matched slightly more often than previously. But increased matching of the previously non-dominant attribute was clear for the older subjects, many of whom now matched colour as often as they did shape.

There is, then, some evidence in the literature to support the present finding that type of task can affect which stimulus attributes are most and least often confused, but the nature of the stimulus set and of the transformations applied to construct the distinguishing levels of an attribute are also clearly important. As already indicated, the range of discriminanda and methodologies to be found in reports on children's performance in matching, classification and discrimination tasks is enormous. The suggestion here is that caution should be exercised in the interpretation of experimental results and comparisons across different studies.

At any rate, the objective established for the first part of the present study has now been attained: the contradictory findings of Taylor and Wales are accounted for. In decisions about the sameness of these schematic faces, the children's responses did indeed display a highly systematic ordering among the attributes. Which particular attribute(s) will be correctly matched and which confused is governed not simply by whether or not there is a frame around the face nor simply by the type of task. It depends on a rather complex interaction between the framed/unframed factor and methodological variables, each alone being a necessary but not sufficient condition to explain the differential response distributions reported by Taylor and Wales.

PART 3

DETECTION OR SELECTION OF STIMULUS ATTRIBUTES?

I. MANIPULATION OF OBJECTIVE VISUAL SALIENCY

The experiments in the preceding section have ascertained that both task context and presence/absence of an external visual frame affected children's decisions about sameness among the first set of schematic faces, altering the distributions of matching responses across the three stimulus attributes. Some obvious further directions for investigating the perceptual-cognitive factors at work here would have been to test the Framed-Unframed effect with new discriminanda (for instance, geometric forms) or to extend the methodological inquiry to cover, say, serial stimulus presentation, recognition, or discrimination-learning techniques. However, a fresh line of study was opened.

What prompted the shift of emphasis was a concern that in the initial experiments the three stimulus attributes and their respective variants were, in purely physical terms, structured quite differently from each other. For one thing, the Shape and Eyes variants preserved the basic outline depicting those attributes, while it was the Mouth outline itself that varied. The two Shape ellipses differed in <u>orientation</u>; Eyes varied by <u>colour</u> contained within the circle outline; and Mouth, by a <u>line-to-curve</u> type transformation. In terms of the amount of the stimulus card occupied, moreover, Shape could be regarded as a "large" feature that was perhaps more noticeable than the relatively smaller Eyes and Mouth.

Might it not then be that objective visual saliency differed among the three attributes and their variants -- that they were not equally physically discriminable (and that, correspondingly, some entire faces were more physically salient than others)? If so, this clearly might affect which attribute is most often correctly matched or confused -- and might interact with the already complex frame x task interaction.

There are essentially two ways of tackling the question of objective saliency. One involves experimental manipulations to vary the degree of discriminability of the stimulus attributes (e.g., by varying their relative sizes) and also both the degree and type of transformation employed in the construction of distinguishing levels within attributes. The other involves equating saliency across attributes and levels by equating their physical properties.

With the present, complex stimulus forms, to follow the first method would be a formidable endeavour if adequate controls were to be incorporated. Although work has been done on physical stimulus discriminability, type of transformation, etc., it is hard to relate to the present situation. Not too surprisingly, heightening discriminability by increasing the degree of physical difference between attributes and their variants leads to better matching or classification performance, even for adults (Imai & Garner, 1965; Garner & Felfoldy, 1970).

With regard to type of physical difference, some work has already been mentioned in the methodological discussion at the end of the preceding section. For instance, the research using letter-like forms and their transformations as stimuli (e.g., Gibson et al., 1962; Gibson, 1965, 1971; Pick, 1965) clearly shows that children can discriminate certain types of variant from the standard more readily than other types.

Schaller and Harris (1974) extended the Gibson work to examine the effect of degree of transformation as well as kind. Thus, for the line-to-curve variants, a transformation of three lines to curves was presumed to be objectively more discriminable than a transformation of two lines to curves; or, for perspective variants, increasing the angle of tilt was taken to increase objective discriminability. Again as might be expected, children's confusion errors became fewer as degree of objective discriminability increased, and the improvement with age was more marked for more saliently different figures.

In this kind of work, however, comparison <u>among stimulus attributes</u> does not enter, since it is the whole stimulus figure, as a global entity, that is treated as transformed. And as Schaller and Harris (1974) point out, it is impossible here to compare degree of discriminability <u>across</u> types of transformation. Returning to the schematic faces employed in the first experiments of the present series, suppose we double the diameter of the Eyes circles. Are they now twice as discriminable as before? And how does this change compare, physically, with doubling the length of the Mouth line?

A study by Hess and Pick (1974) comes closer to the present situation but still falls short of meeting its complexities. Their stimuli were also schematic faces, consisting of two relevant attributes: eyes were ellipses of varying degrees, and mouth was an upward curve (as in a "smile") varying in radius from nearly a straight line to a semicircle, while the circular head-shape was held constant and irrelevant. These authors attempted to equate discriminability between eyes and mouth variants as follows.

First, children were presented with eyes forms and mouth forms separately and out of the face context, for pair-comparison. Frequency of confusion errors was taken as a measure of the discriminability of a given pair of eyes variants or of a given pair of mouths. Eyes and mouth pairs were thus independently rated over their respective ranges of variation from poorly to highly discriminable. This was intended to yield a way of comparing and equating eyes and mouth for degree of psychophysical discriminability.

The features were then combined to make up sets of schematic faces, again to be judged in pair-comparison. For half the face pairs, mouth varied but eyes were constant; for the rest, eyes varied and mouth was fixed. The error patterns showed that most children matched by eyes. Predictably, far more errors occurred on pairs previously rated as poorly discriminable than on highly discriminable pairs. But even with highly discriminable features, mouth was confused significantly more often than eyes. Hess and Pick concluded: "... the evidence suggests that changes in generally positive-appearing [i.e., curved up, 'smiling'] mouths are difficult to attend to while <u>equivalent</u> changes in eyes are somewhat more noticeable" (1974, p. 1154, emphasis added).

Such a bold claim seems unwarranted on the basis of their study. There is an obvious problem: we do not know that the changes in eyes were <u>equivalent</u> to the changes in mouths. From the first part of their study, we know only that certain eyes variants were more discriminable than other <u>eyes</u> pairs, and that certain mouth variants were more discriminable than other <u>mouths</u>. We cannot tell if the most (least) discriminable eyes pairs were <u>equally</u> easy (difficult) as the most (least) discriminable mouth pairs. It could perhaps be argued that eyes differences were more salient than mouth since (a) there were <u>two</u> eyes but only one mouth, and (b) the ellipses of the eyes could be viewed as composed of <u>two</u> curves, one upper (curved down) and one lower (curved up), enclosing a solid area, whereas mouth was just one one curve (always upwards). While Hess and Pick's speculation that the mouths were less differentiable because they tended to be classed together as all "smiling" is appealing (and will be returned to later), they failed in their intention to equate discriminability across eyes and mouth, and their data do not justify their broad suggestion "that eyes are more important sources of information for perception of faces than are mouths" (p. 1151, Abstract).

This study illustrates some of the difficulties of experimental designs under which discriminability effects are examined by attempting systematically to manipulate objective visual saliency of stimulus attributes. There is also a problem of interpretation in this sort of research, for it is often simply assumed that when a child correctly matches a particular feature he is <u>deliberately attending</u> to it rather than to another. The work of Eleanor Gibson and her colleagues generally begs the question of whether errors arise because the child intentionally selects certain stimulus features as "criterial" and ignores others, or because certain types of transformation are more physically detectable than others. It seems to be taken for granted that correct performance on a particular kind of transformation reflects its "distinctiveness" as a criterion of difference. Taylor and Wales (1970) also merely presumed that their observed uneven distributions of confusion errors among stimulus attributes were due to the subjects' deliberate weighting of one attribute over another.

Hale and Lipps (1974) to some extent avoided this trap by their comparison between triads and component-selection tasks, since in the latter the child at least has the opportunity to show whether or not he has noticed and can use the non-dominant attribute when appropriate (see pp. 103-104 above). But in other cases it can be difficult to assess the effect of manipulations of stimulus discriminability, because of potential interactions between <u>objective</u> and <u>subjective</u> saliency. The Hess and Pick (1974) study nicely demonstrates part of this problem -- it is one of circularity. That is, in the end, the only measure of salience and discriminability that we have in this sort of research is a psychological rather than a physical one; it comes from the subjects' behaviour. Schaller and Harris state the dilemma succinctly: "The only way to define equivalence across transformations is on the empirical basis of responses to the transformations. To subsequently make assertions about performance on the basis of equivalence thus defined would be tautological" (1974, p. 230, footnote 4). Yet this is indeed what has been done in some of the so-called psychophysical studies of similarity.

The problem, while evident, is yet enormously difficult to overcome. Subjects' responses are influenced by both subjective and objective perceptual-cognitive factors, but how do we separate out the two? As Gregson notes, psychometric theory generally fails to represent the full complexity of similarity-judgment behaviours because it lacks the capacity to handle the "nonlinear combinations of measures of physical attributes" that such psychological behaviours incorporate (1975, pp. 196-197). In other words, people -- be they children or adults -- do not always behave in ways that reflect simply the physical composition of the stimuli. They may subjectively assign differential importance, or "weights", to the various stimulus attributes, whose physical properties may or may not bear on the weightings (see also Goodman, 1970).

Evidence for the operation of such subjective factors in practice comes from, e.g., Reed (1970). Using stimulus categories consisting of exemplars from Brunswik's set of schematic faces, Reed examined how adults go about classifying the faces, and tested various proposed models of classification strategy. Models incorporating a weighting component gave the best fit to subjects' actual responses, with the Weighted

Prototype model generally superior to the Weighted Average Distance model. That is, when matching Brunswik faces, it appeared that most subjects formed a notion of "the prototypical category member" and classified test stimuli according to their match with the abstract prototype. Thus subjects may use stimulus aspects that are not structurally specified, such as relationships among stimulus attributes, goodness of form, etc.

Tversky (1977) extended this line of research both theoretically and empirically in his cogent review of a series of studies on similarity-dissimilarity judgments by adults and of various possible models to account for them. He, too, argued that stimuli cannot always be defined by their physical components -- "more global properties (e.g., symmetry, connectedness) should also be introduced" (1977, p. 331). In Tversky's theoretical framework (called the contrast model -- see further Gati and Tversky, 1982), both features common to members of a stimulus set and distinguishing features unique to a member are important; perceived similarity is conceived as resting on a weighted difference between common and distinguishing features. This implies that within a given stimulus set there are several possible ways of defining similarity relations among members -- exactly how people partition a set is of course what is empirically determined.

Incorporated into the model is a measure of relative salience among stimulus features. By "salience", Tversky means a compound of both objectively- and subjectively-defined prominence, resting on such factors as intensity, frequency, familiarity, good form, and informational content. The objective (or, as Tversky calls them, intensive) factors are those "that increase intensity or signal-to-noise ratio", and are relatively context-independent. The subjective, or "diagnostic factors refer to the classificatory significance of the features" (p. 342).

The classificatory, or diagnostic, value of a feature is highly context-dependent, as Tversky nicely demonstrated. By the simple addition or substitution of a single member in a stimulus set (including a set of schematic faces), adult subjects given the task of dividing the set into two groups were induced to alter their basis of classification, and also correspondingly to alter their judgments of similarity among the members.

So far in the present study, the discussion has avoided the issue of whether a confused stimulus attribute is confused because it is not detected, or because it is deliberately ignored. The former case has to do with *objective* saliency: confusions arise when one attribute is less noticeable than the others and is accidentally overlooked, or when the physical separation of its distinguishing levels is of a size or character that is not readily discriminable by young children. In the latter case, *subjective* saliency comes into play: attributes are weighted as to their importance for the task in hand, such that certain attributes become criterial and are correctly matched, while others are treated as irrelevant and are confused (or matched just by chance).

We have already seen that, with the schematic faces used in the first three experiments, the observed patterns of differential correct matching among the attributes could, by virtue of their different physical properties, be due to differential objective saliencies. But the results, sketched above, of Reed (1972) and Tversky (1977) with other schematic face forms suggest that adults do rank features for their importance in similarity judgments, so this possibility cannot be excluded for children. The question needs to be decided empirically.

In view of the confounding of subjective and objective saliency effects in responses, and of the difficulties inherent in any experimental attempt to <u>vary</u> objective saliency, it was decided to follow the alternative approach: to hold objective saliency constant by <u>equating</u> the physical properties of the stimulus attributes. If all attributes are structurally identical and their distinguishing levels are all equally physically discriminable, and if the children then still match one particular attribute and confuse another, we can be more confident that their confusion errors are due to a deliberate weighting of one attribute over another.

Part 3 is devoted to this topic. A single set of schematic faces, now with newly defined attributes, is used over its three experiments. Because it is impossible physically to equate shape of head with internal features, Shape is now an irrelevant attribute; it is simply a constant circle throughout. The surrounding frame is therefore also dispensed with. It is only the internal features that are of interest now. In particular, an almost universal dominance of Mouth over Eyes emerged in the preceding experiments, which requires examination, especially in the light of conflicting evidence from other sources. With a few exceptions (e.g., Ahrens, 1954), there is general consensus that, for infants and children, the upper, eyes part of the face (realistic or schematic) is more effective at eliciting gaze, smiling, correct matches, etc., than is the lower, mouth part (e.g., Fantz, 1958, 1961, 1963, 1966; Brooks & Goldstein, 1963; Goldstein & Mackenberg, 1966; Hess & Pick, 1974; Langdell, 1978; Freeman, pers. comm., 1978 [3]). This suggests that the failure to match Eyes in the present study may have been an artifact of stimulus construction. A third internal feature, Nose, is now added to serve as a control -- and because several subjects had remarked on its former absence ("It's not like a real face")!

All three experiments in this section employ the pair-comparison format, which was preferred to matching-from-sample for greater ease of data analysis and greater clarity of response distributions and matching patterns (although Taylor, 1973, and Taylor & Wales, 1970, would debate that superiority of pair-comparison). It was also considered preferable to the method of triads for, although the results are equally clear and readily analysable, with triads the test session is lengthier (allowing extraneous factors like inattention to influence responses) and completely correct responses are precluded. Pair-comparison admits of both correct "not the same" judgments and correct "same" ones (as well as errors in each), thereby providing further controls.

Since the stimuli and basic task format are constant over the three experiments, they are described first, and only specific variations in procedure will be mentioned under the appropriate experiment heading.

Materials

The figures were drawn in black ink on white card. The three internal features, Mouth (M), Eyes (E) and Nose (N), were all based on the Mouth of the faces in the first stimulus set, used in the preceding experiments. That is, the levels to be distinguished were formed from the same set of straight and curved lines across all attributes. Trial faces were drawn varying the length of line, angle of curvature and position of attributes. The final selection took into account comments offered by adults and children who were informally presented with a range of trial figures, but rested ultimately on the Experimenter's subjective decision. The chosen forms were intended to provide "reasonable" discriminability among levels of attributes along with controls on the structural design of the stimuli, while preserving their "facelike" quality.

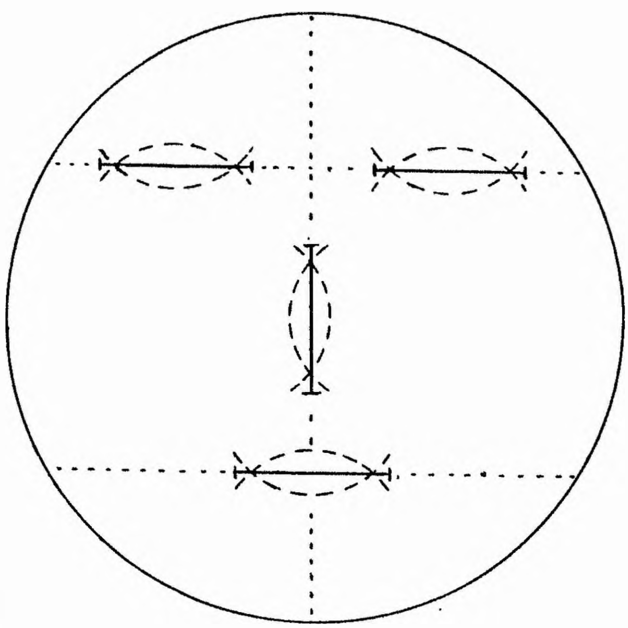
Once their physical properties were established, a template for the attributes was constructed to ensure faithful duplication. This took the form of a small segment cut off a circle of radius 12 mm, such that the chord of the segment was 2 cm long:

Each attribute had three levels: the chord of the segment gave the "neutral" level (a 2-cm straight line) and its arc formed the two curved variants. For Eyes and Mouth, these were curved up and curved down, and for the Nose they were curved left and curved right.

These attributes were assembled to produce a set of schematic faces. Throughout the set, the outline of the head was a fixed circle of diameter 8.5 cm. In order to position the internal features, a basic model (shown in Figure 3, next page) was constructed, as follows. The circle representing the head was drawn, and its vertical diameter entered. Two horizontal chords were added, one cutting the diameter a quarter in from the top, the other a quarter up from the bottom. (The diameter and chords were guidelines and did not actually appear on completed stimuli; they are represented as dotted lines in Figure 3.) On the diameter, the central 2 cms was marked off as a vertical strip designating the location of the Nose. The central 2 cms of the lower chord was marked off for the Mouth. The two Eyes were set along the upper chord, each 2-cm strip being equidistant from the circumference of the circle and its vertical diameter. In Figure 3, solid 2-cm bars show the positions of the attributes in the neutral level. Dashed arcs represent the corresponding curve variants. For these, the arc of the attribute-template was placed over the same area as its neutral counterpart, such that the arc projected equally over either side of the location of the (now imaginary) 2-cm strip.

Figure 3

Construction of the basic stimulus model for placing of Mouth, Eyes and Nose in the second set of schematic faces (see text for explanation)

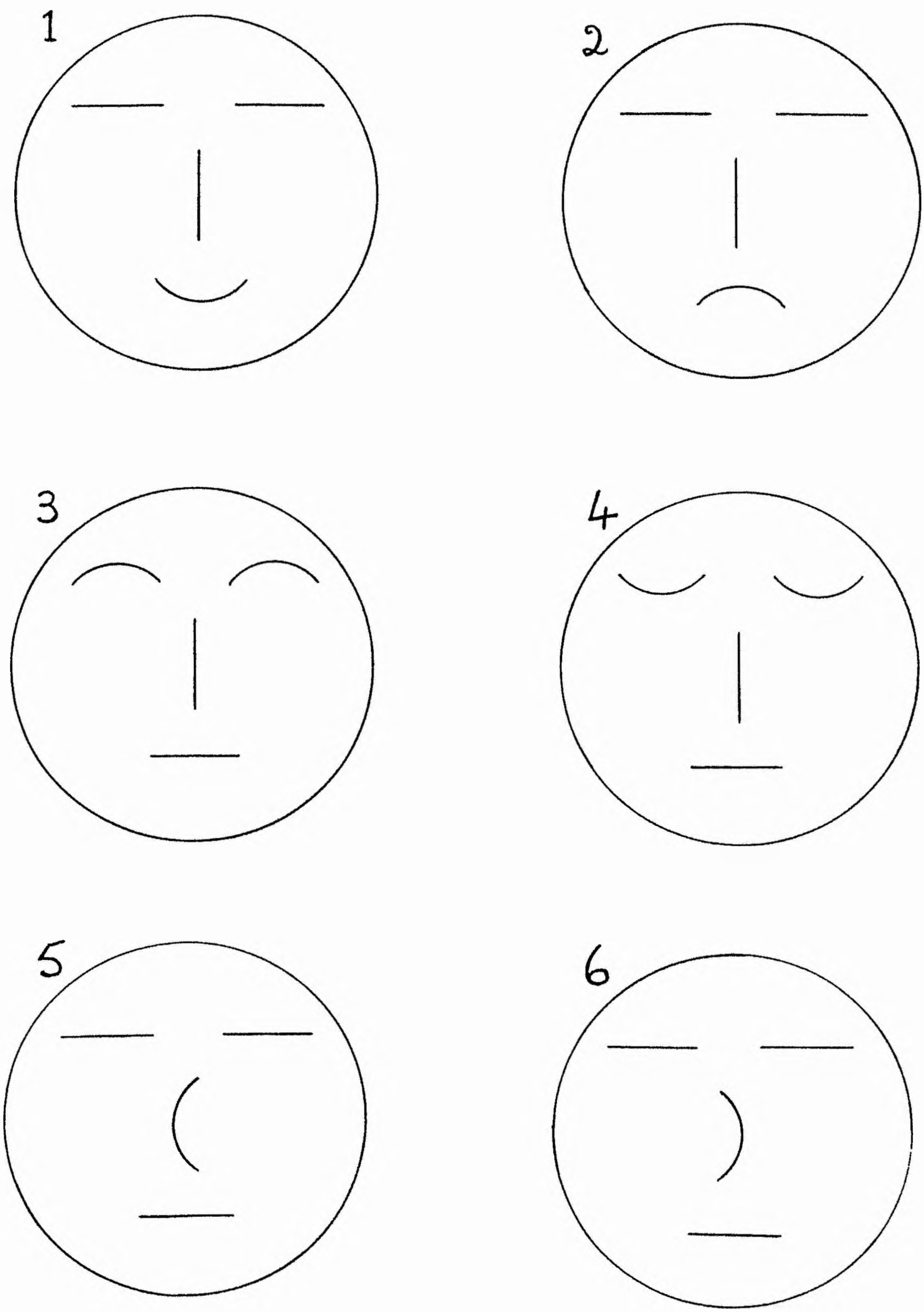


Combining the features in all permutations (three attributes each with three levels) would produce a total of 27 possible faces -- an enormous stimulus set for a young child to cope with, and unnecessary for present purposes. To determine whether one attribute is noticed or used more than another, a subset of only six of these faces suffices. The actual stimulus set therefore consisted of just the six critical members:

(1) E and N neutral, M curved up	(2) E and N neutral, M curved down
(3) M and N neutral, E curved down	(4) M and N neutral, E curved up
(5) M and E neutral, N curved right	(6) M and E neutral, N curved left.

Figure 4

Second stimulus set -- used in Experiments 4, 5, 6 and 7.
(The same three variants make up each of the attributes)



These are perhaps best conceptualised as three stimulus pairs: one pair in which Eyes and Nose are identical (neutral) and Mouth differs (curves); one in which Mouth and Nose are identical and Eyes differ; and one in which Mouth and Eyes are identical and Nose differs. This stimulus set is illustrated in scaled-down proportions in Figure 4 on the preceding page, with a number added to each face so that reference can be made to particular members. This diagram also appears in Appendix 1, Figure 3. Figure 4 of Appendix 1 shows the actual size of the stimuli. Replicas of these six faces were made so that identically-matching stimulus pairs could be produced. Each face was placed centrally on an 11.5-cm square white card (with no surrounding frame) which was then covered in washable transparent plastic film.

For preliminary practice trials, the triangles (two on white cards and two on pink) used in Experiments 2 and 3 were retained, and two new pairs were added. These were coloured figures drawn on white cards and differing in internal detail: two equal-sized squares, one coloured red in the middle and the other blue; and two identical circles outlined in red, both with a long blue chord cutting across the lower right, but in the upper left one had a short blue chord while the other had a blue zig-zag line.

Stimulus presentation

The task was to be the conventional pair-comparison method similar to that employed in Experiment 3 above. Were each of the six stimuli in the present set to be paired both with itself and with every other member, a total of 21 trials would obtain: six identical pairs and 15 distinct ones. For the purpose of determining whether or not children respond differentially to Mouth, Eyes and Nose, however, the test really rests on just three of those trials. These critical pairs have already been mentioned, in the description of the composition of the stimulus set (see foot of p. 117 and Figure 4). They are: faces 1 and 2, with only the Mouth differing; faces 3 and 4, with only the Eyes differing; and faces 5 and 6, with only the Nose differing. Since they involve differences of just one attribute at a time, these pairs should readily show whether or not a particular attribute predominates in the children's responses. A "strict Mouth-matcher", for example, will correctly judge faces 1 and 2 (M different) to be not the same, but will wrongly call pairs 34 and 56 (M identical) the same. An individual child may of course make an "accidental" slip on the odd occasion (e.g., through temporary lapse of attention); but if an error occurs on the same trial consistently across a number of children, it is clear that it is not the product of any such random or chance factor. All other unidentical pairs differ in <u>two</u> attributes (i.e., ME, or MN, or EN). Since, in the present study, two-attribute differences have been associated with comparatively few wrong "same" answers, these pairs are of less value than the single-difference ones. (Notice that no complete mismatches -- pairs differing in all three attributes -- are possible with this stimulus set.)

On the above grounds, and reasoning that with young children it is best to keep the test session brief, not all of the 21 possible pairs were used. Every subject received all six of the identical-match pairs, and all three of the critical pairs differing in a single attribute. Of the remaining 12 pairs (differing by two attributes), each child received only three. Thus the session contained 12 test trials in all, six identical pairs (expecting the answer "same", for correct performance), and six unidentical ones (expecting the answer "not the same"). This arrangement had the advantage of counteracting the usual pair-comparison problem whereby negative responses are expected more often than positive, possibly resulting in a bias towards false-positive answers in the more typical version of the task employing all possible unidentical pairs.

Response sheets were drawn up, one per subject, listing the 12 trials. The order of trials and positions of stimuli within the pairs were randomised across subjects. The three pairs differing in two attributes varied across subjects, taken at random from the 12 such possible pairs.

A further stimulus presentation variable was introduced in the three experiments in this section. Within each attribute, the two curved variants constituted mirror-image reversals of each other -- up-down reversals for Mouth and Eyes, and left-right reversal for Nose. We saw earlier (pp. 8-9) that infants and young children frequently fail to discriminate reversals that are presented side by side as true left-right mirror images, but are reasonably successful when those reversals are aligned one above the other so that they no longer form mirror images. Similarly, reversals that are up-down mirror images displayed in up-down alignment tend to be confused more often than when they are shown side by side.

Given the physical structure of the present set of schematic faces, it is possible that such factors might intrude to interact with, or even override, any effects of responding differentially to the various attributes themselves. If the faces were presented side by side, for instance, one might find the Nose (left-right reversal) to be confused more often than Mouth or Eyes simply because of the mirror-image x stimulus-alignment arrangement alone.

For this reason, two conditions of stimulus presentation were applied in the pair-comparison tasks of Experiments 4, 5 and 6: the face pairs were displayed either in left-right (L-R) alignment or in up-down (U-D) alignment. In each experiment, half of the subjects received the L-R display on all trials, and the other half received the U-D display throughout. Besides serving as a means of counteracting mirror-image effects that might mask or artificially exaggerate differential attribute selection, this design allows one to check on the occurrence of these mirror-image effects themselves. If they are of sufficient magnitude to override any differential detection or use of attributes, then the L-R condition will, as noted, produce more confusions of Nose than of Mouth or Eyes, while the U-D condition will produce more confusions of Mouth and Eyes than of Nose.

Procedure

As before, children were taken individually to "look at some pictures". Subjects in these three experiments came from two kinds of sources: playgroup and nursery class. In the playgroups, a corner of the room containing bench or chairs and table was sectioned off. Nursery children either were tested in a similarly partitioned-off corner of the classroom, or were taken to a vacant staff-room equipped with low chairs and table.

The pair-comparison task was conducted much as in Experiment 3 above. Child and Experimenter sat side by side. Stimulus pairs, in the order and positions previously determined and given on the individual response sheet, were laid on the table in front of the child, the nearer sides of the stimulus cards touching each other. The Experimenter asked, "Are these two (pictures) the same or not the same?" on each trial, and recorded the child's response. In Experiments 4 and 5, the child was not explicitly asked on test trials to explain how the pictures were the same or different, but any spontaneously-offered comments were recorded. (The procedure of Experiment 6 differed in this and following respects and will be described later.) The test trials of Experiments 4 and 5 were preceded by five trials with the practice stimuli -- three with the triangles and two with the coloured figures, in random order. During practice, unlike in the test itself, the child was urged to describe the two stimuli and to indicate specifically points of identity and difference between them. The Experimenter intervened where necessary during practice to ensure that the child understood the task, but gave no feedback save general verbal encouragement during test trials. The session in Experiments 4 and 5 lasted for about 10 minutes.

EXPERIMENT 4

Conventional pair-comparison

Subjects

Forty children were tested, all performing satisfactorily on practice trials and all completing the test session. Nineteen came from a playgroup and 21 from two nursery classes in St. Andrews, spanning a variety of home backgrounds. They were assigned to two groups of 20, one to receive the Left-Right stimulus presentation condition, the other the Up-Down condition. The two groups contained equal numbers of boys and girls, and age was balanced as far as possible. In the L-R Group, ages ranged from 3;4 to 5;2 years with a mean of 4;6; in the U-D Group, the range was 3;5 to 5;2, mean 4;7 years.

The experimental procedure was exactly as described above.

Expectations

This experiment was designed to determine whether or not the children would respond differentially to Mouth, Eyes and Nose now that their physical properties, and hence objective visual saliency, were equated -- and if so, which of the attributes would predominate in correct matches and which would be the most often confused. The principal question was thus open.

In reality, however, specific predictions could be considered possible, for the statement that the three attributes were equally physically salient is not strictly true. The reason is that while Mouth and Nose were depicted by a single line, be it straight or curved, Eyes were represented by two lines. Thus, it could be argued, in the critical test pair for Eyes (faces 3 and 4), there were two distinguishing features, whereas the critical Mouth and Nose pairs (12 and 56, respectively) differed "less". Moreover, the potentially confounding effects of mirror-image reversals should also be taken into account.

What to expect was therefore unclear. The main possibilities considered were:

(1) If objective saliency is the most important contributor to differential distributions of correct judgments and confusions among the attributes, then judgments based on Eyes will predominate and Mouth and Nose will be more often confused.

(2) If mirror-image reversal effects play a major role, Nose will be confused in the L-R condition but will be correctly matched in the U-D condition; and vice versa for Mouth and Eyes. Further, if objective saliency is also important, then independently of the mirror-image effect, Eyes should be consistently more often correctly discriminated than Mouth under both L-R and U-D alignment.

(3) If children's judgments of sameness involve a large degree of subjective weighting of attributes, then one attribute (not initially specified but to be empirically determined) will predominate irrespective of whether the stimuli are presented in L-R or U-D relation. But if the dominant attribute turns out to be Eyes, it will be unclear whether this is a result of objective saliency or subjective weighting factors and further tests will be required.

RESULTS

First, consider the "control" trials that serve as a check on the general level of performance and understanding. These were of two types: those in which the pair was identical and should be judged as the same (six trials per subject), and those in which the pair differed by two attributes and should be judged as not the same (three trials per subject).

Out of the sample of 40, only two boys wrongly called an identical pair different. From the L-R Group, one of 4;7 years made errors on faces 1 and 4 paired with themselves; and one of 3;5 years in the U-D Group made an error on pair 55.* Since these two children correctly judged both the remaining identical pairs and the pairs differing in two attributes, their errors were regarded as insignificant.

As to the two-attribute differences, errors were similarly rare. Six subjects made a total of seven errors in all (out of 120 of this kind of error possible over the entire sample). Two were contributed by

---

* The latter child spontaneously commented "they're funny" on this trial. After he had completed his 12 trials, the Experimenter re-presented pair 55; he still insisted they were not the same "because they're funny". It was as if he meant that face 5 was in some way peculiar in relation to the rest of the stimulus set. This will be returned to later (in the corresponding results of Experiment 5, and in Part 4).

younger boys in the L-R Group, and the rest by U-D subjects of both sexes and a range of ages. No errors occurred for any pair differing in MN; four were on EN differences and three on ME differences -- the latter all involving the curved-up ("smiling") Mouth (pairs 13 and 14). But the total number of errors was too few to allow conclusions about any systematic bias in their distribution. As expected, then, the children did generally answer correctly on trials involving two points of difference, and the limitation of these trials to three per subject instead of all 12 seems to have been justified.

Six subjects were entirely correct on their 12 trials. All belonged to the L-R Group, which suggests that horizontal stimulus alignment may have been somehow "easier" for the children to cope with than the vertical relation.* Entirely correct performance was not related to age, but only one of these six children was male -- and he was among the oldest subjects (5;1 years). Coupling this with the fact that six out of the eight subjects who made errors on identical-match and two-attribute difference trials were boys, the suggestion is that girls tended to perform better than boys. This is the first hint in the present series of experiments of any noticeable sex-related difference in responses. But the numbers here were small and sex differences did not appear in later experiments, so it is probably a chance result.

---

* Possibly the mirror-image reversal effect may have played some part here, although the left-right reversal of Nose should then have been highly confusable under Left-Right stimulus display, more so than under Up-Down display. The results of Experiment 5 also argue against any significant mirror-image effect. Note, additionally, that most of these children were too young to have had much experience of reading and writing, i.e, specific training in left-to-right visual processing.

Errors were considerably more frequent on the three critical trials involving a single-attribute distinction: pair 12 differing in M, pair 34 differing in E, and pair 56 differing in N. On these trials, each subject was assigned a score of 1 for a correct "not the same" answer and a score of 0 for a wrong "same" answer.

Overall, the Mouth distinction was successfully discerned by 77% of the children, Eyes by 47% and Nose by 25%. The corresponding figures for the two stimulus display conditions separately (n = 20 per Group) are: L-R condition -- 90% correct on Mouth, 65% on Eyes and 35% on Nose; U-D condition -- 65% correct on Mouth, 30% on Eyes and 15% on Nose. Thus the pattern of differential responding is essentially unaltered between the two display conditions, although the U-D Group made a greater number of errors in general. This agrees with the earlier comment, based on the finding of no entirely correct sessions under U-D alignment, that comparison of stimuli in U-D relation seems to have been generally harder.

Cochran $\underline{Q}$ tests (two-tailed) were conducted on these data, the alternative to the null hypothesis being non-directional since no prediction was formulated in advance as to which, if any, attribute would predominate. The resulting values confirmed that the probabilities of correct "not the same" judgments differed significantly among the three attributes (i.e., among the three critical face pairs). This was true for both the pooled scores of the entire sample of 40 subjects ($\underline{Q}$ = 23.786, $p < 0.001$) and for each stimulus display condition separately (L-R Group, $\underline{Q}$ = 14.0, $p < 0.001$; U-D Group, $\underline{Q}$ = 10.533, $p < 0.01$) (all with df = 2).

To examine more precisely where the differences resided, Binomial tests were applied between pairs of attributes. For these, the data were expressed in the form of the number of subjects who were correct on one attribute but not the second (e.g., who scored M > E) compared against the number who were correct on the second but not the first (E > M). All tests were again two-tailed. For the pooled responses of the entire sample, all values reached significance: for Mouth vs Eyes, $p = 0.002$ ($N = 14$, $x = 1$); for Mouth vs Nose, $p < 0.002$ ($N = 25$, $x = 2$); and for Eyes vs Nose, $p = 0.05$ ($N = 17$, $x = 4$). Considering each stimulus condition separately, the following reached significance: Mouth vs Eyes for U-D Group only ($p = 0.016$, $N = 7$, $x = 0$); Mouth vs Nose for both groups (L-R Group, $p < 0.002$, $N = 11$, $x = 0$; U-D Group, $p = 0.012$, $N = 14$, $x = 2$).

The data on which the above computations were based are given in Table 22. Here, the numbers of correct responses per attribute are shown for subjects divided into age groups as well as by stimulus display condition. Subjects had been assigned to age groups at the outset (in contrast with the preceding experiments, where age divisions were made on a post-hoc basis). Each subgroup contained five boys and five girls:

L-R condition, younger -- ages 3;4 to 4;9, mean 4;1 years
older -- ages 4;8 to 5;2, mean 4;11
U-D condition, younger -- ages 3;5 to 4;8, mean 4;2
older -- ages 4;8 to 5;2, mean 4;11.

Overall, the older children attained the greater number of correct responses on the Mouth-differing pair. But the two age groups were about equal on Eyes responses, and the older children actually made fewer correct judgments of the Nose distinction than did the younger. In contrast with the preceding experiments, then, there was here no clear general pattern of improvement with age. (There were no apparent sex-related trends, so the earlier suggestion from the entirely correct subjects and the two-attributes-different trials that girls might perform better than boys was not borne out.)

TABLE 22 (Experiment 4)

Number of correct "not the same" judgments of critical pairs differing in one attribute, by stimulus alignment condition and age

	Younger Subjects			Older Subjects		
stimulus pair and distinguishing feature:	12 / M	34 / E	56 / N	12 / M	34 / E	56 / N
L-R Group (n = 10 per cell)	8	7	4	10	6	3
U-D Group (n = 10 per cell)	5	2	3	8	4	0
All subjects (n = 20)	13	9	7	18	10	3
% correct judgments	65%	45%	35%	90%	50%	15%

TABLE 23 (Experiment 4)

Distribution of 28* subjects according to which of the critical pairs (attributes) they correctly judged to be not the same

	STIMULUS PAIR(S) & ATTRIBUTE(S) CORRECTLY DISTINGUISHED					
	12 M	34 E	56 N	12&34 M&E	12&56 M&N	34&56 E&N
L-R Group (n = 13)	5	1	0	6	1	0
U-D Group (n = 15)	6	0	2	6	1	0
All subjects (n = 28)	11	1	2	12	2	0

*The remaining 12 subjects of the sample are omitted as they were either correct on all three or wrong on all three critical pairs.

The pattern of Mouth correct and Eyes/Nose confused was also striking when individual subjects rather than pooled scores were considered. Of the sample of 40 children, six correctly distinguished all three critical pairs and another six failed on them all. Thus 28 children were correct on one or two but not all three critical trials. Table 23 shows the distribution of these 28 subjects according to which of the critical pairs they judged correctly, i.e., which of the attributes they succeeded in discriminating. It is clear that Nose was distinguished by very few of the subjects and Mouth by most -- and that a child did not judge Eyes correctly unless he was also correct on Mouth. This distribution obtained irrespective of stimulus alignment condition.

Summary and conclusions

It is quite evident that the results follow the pattern described as the third of the outcomes considered possible (see pp. 124-125). The children did respond differentially to the three attributes; and the same pattern, in terms of which particular attribute was most often correctly judged and which most often confused, obtained across both conditions of stimulus display. That is, regardless of stimulus alignment, Mouth-based decisions predominated and Nose was the most often confused.

This finding counters both the objective saliency and the mirror-image reversal hypotheses which would predict, respectively, either that Eyes predominate or that the distribution of responses among the three attributes differ between the two stimulus conditions. It appears likely, then, that the children subjectively assigned differential weights to the attributes, treating Mouth as most important and Nose as least important to the question of whether the faces were the same or not. We cannot yet state this firmly, however, for it could be argued that the way the subjects visually inspected the stimuli might have been at least partly responsible for the observed pattern of correct judgments and confusion errors: perhaps the children looked first to the <u>bottom</u> of the stimulus figures (Mouth) and last, if at all, to the central detail (Nose). (A tendency to look at the bottom first could be invoked to explain the general predominance of Mouth over Eyes matches in the first three experiments, too.) The following experiment was designed to examine this possibility.

EXPERIMENT 5

Pair-comparison with inverted stimuli

There was a simple way to test whether the predominance of Mouth in the correct judgments of Experiment 4 was due to a tendency to focus first or mainly on the lower part of the stimulus figure: repeat the same task with the same stimuli but now upside-down. If the location-of-gaze hypothesis is true, then whatever lies in the bottom part of the figure should still be the most often correctly discriminated attribute -- Eyes, in the case of inversion. This may be too simplistic a view, however, for potential complicating factors arise in the use of inverted stimuli, and in the relationship between visual inspection and location within the stimulus figure of the distinguishing feature.

While young children in free play appear easily to identify pictures viewed in unconventional orientations, there is evidence to the contrary from controlled laboratory investigations. One of the earliest rigorous examinations of children's perception of forms in canonical and inverted orientations seems to have been that by Hunton (1955), using a range of stimuli from realistic (photographs, magazine illustrations) to schematic (simple line drawings). Her subjects were required to signal stimulus identification by verbally describing what was depicted. Following Hunton, Ghent (1960) presented children with realistic pictures of a variety of items in upright, 180^{0} and 90^{0} orientations, in a kind of recognition-cum-serial-matching-from-sample task. The child was briefly shown a picture and then had to either name what he had seen or point to an identical match from an array. Brooks and Goldstein (1963; see also Goldstein & Chance, 1964) gave children a series of photographs of faces of their classmates in upright orientation, which they had to learn to recognise. Later they were asked to identify those previously seen from an array of inverted photographs of faces.

Despite the variety of stimuli and tasks over these studies, there is general consensus among the findings. Older subjects -- that is, at least over age five years -- perform about equally well with both inverted and conventionally-orientated stimuli. Children younger than five years are less successful with inverted than with upright forms. It appears, then, that "something" happens in the immediately pre-school period that prompts facilitation of handling of stimuli in unconventional orientations. Gestaltist or experientialist accounts fail here, for they would imply that as the child is repeatedly exposed over time to objects in one particular orientation, he would come to recognise the most familiar orientation the most readily; i.e., contrary to what actually occurs, older children would be expected to have more difficulty with inverted stimuli than younger.

Ghent (1960) speculated on an alternative possible account, which she proceeded to investigate systematically (Ghent et al., 1960; Ghent, 1961; Ghent & Bernstein, 1961; Braine, 1965; Barroso & Braine, 1974;

see also Wohlwill & Wiener, 1964). The proposal, supported by her subsequent results, was that there is an interaction between success in stimulus identification/recognition and starting-point and direction of visual scanning. She found that subjects, regardless of age, tended to scan stimulus forms in a downwards direction. But while those over age five years started the scanning at the <u>top</u> of the figure, under-fives started at the "focal part" of the figure (which may not coincide with its topmost point) and proceeded downwards from there. Young children may thus overlook the upper portion of the stimulus in certain circumstances. This was specifically tested (Braine, 1965) using stimuli with a focal part at one end and a distinguishing feature at the opposite end, the distinguishing feature being the sole point of difference among members of a stimulus set. Each figure was briefly exposed, once in the upright position with the focal part at the top and distinguishing feature at the bottom, and once inverted. After each exposure, the subject had to indicate from an array the figure just seen. Under-four-year-olds performed worse with inverted stimuli, while the oldest group (4½ to 5 years) performed <u>better</u> with inverted than upright forms. This supported the view that the older children scanned from the top and were helped by having the distinguishing feature located at the top, while the youngest group tended to overlook the distinguishing feature when it was uppermost. (The middle age group, 4 - 4½ years old, who showed no superiority of one or other orientation, were taken to be in a transition phase.)

With regard to the present study, then, the proposal that the children in Experiment 4 might have focussed first on the bottom of the stimulus and then on the top as a <u>general</u> scanning habit seems questionable in the light of the Ghent Braine work. What is possible is that the

Mouth of these schematic faces might, for some reason, have been treated as a "focal part". In that case, the effect of presentation of the faces in inverted orientation would not be clear in advance. If Mouth was indeed a focal part, and if it continues to be when the faces are upside-down, then performance on Mouth should be enhanced. But the common finding noted above that young children are generally less well able to recognise inverted than upright stimuli may complicate the matter. Further, the unconventional orientation may destroy the facelike nature of the stimuli, such that Mouth may no longer serve as a focal part, if it ever did so. (Several adults shown these stimuli upside-down commented that at first glance they looked like just configurations of lines within a circle, and identified them as inverted faces only after further inspection.) Nevertheless, a pair-comparison task with the faces inverted should at least reveal whether or not the bottom part of the stimulus figure is most often correctly discriminated regardless of its content (or regardless of what its content is taken to represent in terms of the present face context).

Subjects

The initial sample consisted of 30 children divided into two groups, one to receive Left-Right stimulus alignment and the other Up-Down alignment. Three were unwilling to complete the task. Of the remainder, 18 came from two nursery classes in St. Andrews and nine from a playgroup in Oxford. The final L-R Group contained seven boys and six girls, aged from 3;7 to 5;2, with a mean of 4;6 years. In the U-D group were eight boys and six girls, aged 3;8 to 5;2 with a mean of 4;8 years. (The slight imbalance between the two groups was due to the three subjects who were dropped.)

Materials and Procedure

These were exactly as for Experiment 4 except that all test stimuli were presented upside-down.

Predictions

For reasons already outlined, several possibilities of outcome could be entertained. The null hypothesis was that there would be no difference among the three attributes in the frequency with which each is correctly judged to be not the same. Such a result might obtain if, for example, comparison were generally more difficult under stimulus inversion than under conventionally-orientated presentation of these schematic faces, and/or if inversion were to disrupt interpretation of the forms as faces. Possible alternatives were:

(1) Location-of-gaze hypothesis. The question is whether the response bias observed in the preceding experiment could be attributable to a focus of visual attention primarily on the bottom of the stimulus. The main present aim is thus to determine whether or not the children exhibit a general tendency to distinguish detail in the lower part of the stimulus correctly more often than that in the upper part regardless of what that detail is or "signifies". The hypothesis here is that, with the faces inverted, Eyes (now at the bottom) will be the predominantly correct attribute.

The remaining alternatives were carried over from the previous experiment:

(2) Subjective weighting hypothesis. If the earlier finding of Mouth-dominance in correct judgments was due rather to a deliberately selective attention to this attribute as the "most important" -- if it served as a kind of "focal feature", to adapt Ghent Braine's term -- then Mouth is expected again to predominate (assuming that facelike interpretation is sufficiently preserved under stimulus inversion).

(3) Objective saliency hypothesis. Eyes will be most often correctly judged because distinct levels of this attribute are the most noticeable, involving <u>two</u> differing lines.

(4) Mirror-image hypothesis. If mirror-image reversal effects come into play, then the two stimulus conditions will produce dissimilar response patterns such that in L-R presentation Mouth and Eyes will be correct and Nose confused, while in U-D presentation Nose will be correct and Mouth and Eyes confused.

Although neither objective saliency nor mirror-image effects were in evidence in Experiment 4, they cannot be completely ruled out in the present situation, since stimulus inversion could affect the visual comparison process to bring them into force. Notice that hypotheses (1) and (3) both predict the same outcome, though for separate reasons. If Eyes turn out to be the dominant attribute, then, further tests will be required to decide between them.

Despite the fact that various outcomes could be regarded as possible, a specific prediction was made in advance: that subjective weighting would operate, favouring Mouth (the second hypothesis). There were two grounds for this supposition. Firstly, although deserving of examination, the location-of-gaze hypothesis was considered an unlikely candidate. The Ghent Braine research on visual scanning and focal parts suggested that children do not look initially at the bottom of the stimulus figure simply because it is the bottom. They may do so if that is where the focal part is located, but otherwise they would tend to look at the top. Secondly, although the children's verbal comments and justifications for their decisions in the preceding experiments had not yet been examined systematically, the intuitive impression from them was that the children were treating Mouth differently from the other attributes in their talk and hence also likely in their decision processes. (Justifications will be discussed more fully in Part 4.)

RESULTS

The various aspects of the results will be presented in roughly the same order as for Experiment 4, beginning with the control trials involving identical matches and two-attribute differences (cf. pp. 125-126).

Identical pairs: Wrong "not the same" judgments. Four subjects in the L-R Group made a total of five errors of this sort, and one in the U-D Group made one error. Although slightly more than occurred in Experiment 4, then, false negative answers were still quite rare.

No sex or age patterns were evident here, but there was one unexpected tendency. In Experiment 4, of the two identical-match errors made, one child was recorded as wrongly calling pair 55 not the same, adding "(because) they're funny" (p. 125, footnote). At the time, this was glossed over as a minor idiosyncracy. However, four of the six identical-match errors in Experiment 5 involved pairs 55 or 66 -- the only members of the stimulus set to have a curved Nose -- and in each case the child made some remark giving the impression that these faces were being treated in a special way because of the Nose: "That one's that way" (pointing to Nose); "There's no '1' on it" ("1" referred to the neutral Nose level, a vertical straight line); "Not the same" (pointing to Nose). Further, one of these children confronted on another trial with pair 15 said not the usual "They're not the same", but "THAT one isn't the same" (pointing to the Nose on face 5 and stressing the word "that"). These comments add weight to the suspicion noted earlier (p. 125, footnote): the children seemed to be saying not that pair 55 or 66 contained two unidentical stimuli, but that faces 5 and 6, with their curved Noses, were strange-looking and different from the rest of the stimulus set. This raises an interesting issue concerning the children's interpretation of the test question "Are these two the same or not the same?", which will be returned to in Part 4.

Two-attribute distinctions: Wrong "same" judgments. Again there were slightly more errors than in Experiment 4 but they were still comparatively rare. In the L-R Group, four subjects made a total of five such errors, and in the U-D Group four subjects made six errors between them. No one particular pair nor configuration of attributes seemed especially to attract errors or correct judgments. There was no age pattern, but only one girl in each subject group made this kind of error. Since no child was wrong on all three two-attributes-distinct trials and all had performed satisfactorily in practice trials, these few errors were considered insignificant.

Entirely correct sessions. Only three subjects managed correct judgments throughout their 12 trials: two boys, aged 4;9 and 4;10 years, from the L-R Group and a girl aged 5;0 from the U-D Group -- all among the oldest subjects.

The above findings taken together suggest that pair-comparison was indeed slightly more difficult in general with the stimuli inverted than with them upright -- but only marginally so. The fact that only older children (approaching five years) attained perfect performance with inverted stimuli also agrees with the reports of previous investigators.

Critical trials: Correct "not the same" judgments. The results on the critical pairs 12 (Mouth), 34 (Eyes) and 56 (Nose) are presented in Table 24, to which the corresponding figures from Experiment 4 have been added for ease of cross-reference. Correct "not the same" responses were

TABLE 24

Percentage of judgments that were correct ("not the same") on critical pairs differing by a single attribute, for Experiment 4 (upright presentation) and Experiment 5 (inverted presentation)

	EXPERIMENT 5 (INVERTED)			EXPERIMENT 4 (UPRIGHT)		
stimulus pair and distinguishing feature:	12 M	34 E	56 N	12 M	34 E	56 N
L-R group*	76.9	30.8	53.9	90.0	65.0	35.0
U-D group*	64.3	14.3	35.7	65.0	30.0	15.0
All subjects*	70.4	22.2	44.4	77.5	47.5	25.0

* L-R groups: Experiment 5, n = 13; Experiment 4, n = 20

U-D groups: Experiment 5, n = 14; Experiment 4, n = 20

All subjects: Experiment 5, N = 27; Experiment 4, N = 40 again assigned a score of 1, and wrong "same" responses a score of 0. As in Experiment 4, the U-D condition produced a greater number of errors in general than did L-R alignment. Despite the slightly lower percentages than previously, it is very clear that Mouth continued to be the attribute most often successfully discriminated (70% correct overall in Experiment 5, as against 44% correct for Nose and 22% for Eyes).

In view of the evidence in the literature, discussed in the introduction to Experiment 5, that very young children have more difficulty than older ones (of around age five) with inverted stimuli, at least in recognition tasks, the data were inspected for age-related trends. The two subject groups were subdivided post hoc:

L-R Group, younger, n = 7, ages 3;7 - 4;6, mean 4;3 years
older, n = 6, ages 4;7 - 5;2, mean 4;10
U-D Group, younger, n = 7, ages 3;8 - 4;8, mean 4;4
older, n = 7, ages 4;9 - 5;2, mean 5;0

The results are presented in Table 25, as percentages of children per subgroup making correct "not the same" judgments. The pattern is rather like that observed in Experiment 4 (see Table 22, p. 129). That is, the older children performed better than the younger on Mouth but were less successful than their juniors on Nose (although, unlike previously, the older children achieved noticeably higher scores for Eyes than did the younger). Upside-down stimulus presentation, in relation to upright, does not seem to have had any general particularly adverse effects for the younger subjects nor facilitatory ones for the older, then. In the subgroups of age x stimulus condition, however, the numbers were so small that no firm statements are warranted.

Little new is gained by looking at individual subjects instead of pooled scores, but for the sake of completeness these figures have been

TABLE 25 (Experiment 5)

Percentage of judgments that were correct on critical pairs differing in one attribute, by stimulus alignment condition and age

	YOUNGER SUBJECTS			OLDER SUBJECTS		
stimulus pair and distinguishing feature:	12 M	34 E	56 N	12 M	34 E	56 N
L-R Group	71.4	14.3	57.1	83.3	50.0	50.0
U-D Group	57.1	0	42.9	71.4	28.6	28.6
All subjects	64.3	7.1	50.0	76.9	38.5	38.5

TABLE 26 (Experiment 5)

Distribution of 18* subjects according to which of the critical pairs (attributes) they correctly judged to be not the same

	STIMULUS PAIR(S) AND ATTRIBUTE(S) CORRECTLY DISTINGUISHED					
	12 / M	34 / E	56 / N	12&34 / M&E	12&56 / M&N	34&56 / E&N
L-R Group (n = 9)	2	0	1	2	4	0
U-D Group (n = 9)	4	0	1	1	3	0
All subjects (n = 18)	6	0	2	3	7	0

*The remaining 9 subjects of the sample are omitted because they were either entirely correct or entirely wrong on all three critical pairs.

included (Table 26; cf. Table 23, p. 130, for Experiment 4). The three subjects who were entirely correct are omitted, as are a further six who made wrong "same" judgments on all three critical trials. Of the remaining 18 children, the majority were correct either on just the Mouth (pair 12) or on Mouth plus one other attribute, usually Nose (pairs 12 and 56). Few were correct on Nose alone, and none on Eyes alone; if the child discriminated one of these attributes, he also usually discriminated Mouth. This parallels the finding of Experiment 4.

The data summarised in the left half of Table 24 were subjected to statistical analyses as in Experiment 4, except that one-tailed significance levels were used since a directional prediction was now being tested (i.e., that Mouth would predominate in correct responses). According to Cochran $\underline{Q}$ tests, the probabilities of correct "not the same" judgments differed significantly among the three attributes: for all subjects ($N = 27$), $\underline{Q} = 14.111$, $p < 0.0005$; for the L-R Group alone ($n = 13$), $\underline{Q} = 6.0$, $p < 0.025$; and for the U-D Group ($n = 14$), $\underline{Q} = 8.222$, $p < 0.01$ (all with $df = 2$). Binomial tests were conducted between pairs of attributes, as before (see foot of p. 127 to top p. 128). Only the following reached significance (at the one-tailed level). Mouth vs Eyes: for all subjects, $p < 0.001$ ($N = 13$, $x = 0$); L-R Group, $p = 0.008$ ($N = 7$, $x = 0$); U-D Group, $p = 0.016$ ($N = 6$, $x = 0$). Mouth vs Nose, only for all subjects pooled: $p = 0.033$ ($N = 11$, $x = 2$).

These tests confirm that the strong response bias towards Mouth persisted when the faces were presented upside-down. From Table 24, however, it is apparent that the results of Experiments 4 and 5 differed with respect to frequency of correct Eyes and Nose judgments. Responses on these two attributes exhibited a near-perfect cross-over effect between the two experiments. Under normal stimulus orientation, Eyes were discerned by about half the subjects and Nose by only a quarter -- and vice versa under stimulus inversion.

This finding could be regarded as congruent with Ghent Braine's theory, summarised in the introduction to Experiment 5, if we assume that Mouth acted as a "focal feature". The rationale is then as follows. When the stimuli were presented upright, the children looked first to Mouth. Since there was nothing below Mouth, most of them reverted next to top-down scanning, jumping up to Eyes and coming to Nose last, if at all (bearing in mind that young children may terminate scanning before all parts of a figure have been covered). Hence the rank order of correctness of the attributes found in Experiment 4: M > E > N. When the stimuli were presented upside-down, the children again looked first to Mouth -- but could now proceed directly downwards, to Nose next and Eyes last. Hence the Experiment 5 order, M > N > E.

While such an explanation is -- in the absence of further testing -- quite speculative, it has appeal, for there is another point of agreement between the present response patterns and those of Braine (1965). Braine's results, recall, suggested that a tendency to start scanning from the top of a figure rather than from the focal part developed as the child approached five years of age. In the present study, the Mouth dominance pointed to primary attention to the focal feature regardless of the subject's age. But there is some indication in the age-group data of Tables 22 and 25 that after looking to Mouth the older children showed a stronger tendency to move on to strict top-down scanning than did the younger. With upright stimulus orientation, only 15% of the older children distinguished Nose where 35% of the younger group succeeded (out of 20 in each age group). Although minor, this pattern conforms with the notion that the older children looked to Nose last, while some of the younger group may have gone to Nose directly after Mouth, scanning upwards from the focal feature. A similar argument could be applied in the case of inverted stimulus presentation (Table 25). There, rising-five-year-olds performed at an equal level on Nose and Eyes (five out of 13 children correct) but only one younger child distinguished the Nose. Perhaps at least some of the older children proceeded systematically downwards from Mouth to encompass both Nose and Eyes, while their juniors tended to be less systematic and switched from the top to the bottom of the figure, missing the central detail?

Summary and conclusions

Mouth has been shown clearly to predominate over Eyes and Nose in correct pair-comparison judgments regardless of whether the faces were presented in conventional or inverted orientation. The Ghent Braine type of explanation fits these data rather well: Mouth could be viewed as a sort of focal feature gaining the child's visual attention first. More speculatively, the differential between Experiments 4 and 5 in order of correctness among the three attributes and the slight age-group differences are in line with Ghent Braine's suggestion of a strengthening top-down scanning strategy with increasing age. In the present situation, however, top-down scanning would have to be regarded as secondary to attending to the focal feature -- perhaps the use of schematic faces rather than geometric forms played some part here.*

---

* See also the earlier caveat on p. 12 and Vurpillot's (1976) criticism of Ghent Braine's work. The (un)likelihood of visual inspection strategy as a significant contributor to the observed response patterns in these experiments will be taken up again in Part 4 (pp. 198-200).

With stimulus inversion, the children evidently did not focus primarily on the bottom of the stimulus (Eyes), nor did they operate on the fact that the Eyes distinction involved two lines of difference. Moreover, the distribution of correct responses over the three attributes remained essentially unchanged across the two conditions of stimulus alignment. In other words, the findings run counter to what could be predicted according to the location-of-gaze, objective saliency, and mirror-image hypotheses. The subjective weighting hypothesis stands unrejected.

The expression "subjective weighting" (of one attribute relative to another) is used here in preference to Ghent Braine's "focal part" terminology. One reason is that to talk of a stimulus as having a focal part could be taken to suggest an all-or-nothing situation: one part of a stimulus is focal and important, and all the rest of the figure is non-focal and not important. With the present schematic faces, it seems more likely that each attribute is being ranked on a scale from more important to less important, as it were.

A second reason is that Ghent Braine seems to imply something mainly visual -- which part of a figure is focal appears to be defined largely by the physical structure of the figure. In the case of the present schematic faces, this seems inappropriate because of the equation of their physical properties across attributes. "Subjective weighting", on the other hand, admits of both visual and cognitive components. That is, one stimulus feature (Mouth, in the present case) is treated as more important than others for the task in hand partly because of physical structure but also because of the subjective significance that an entire stimulus figure or its individual features can have for the perceiver-as-interpreter.

One further point deserves discussion. At the start of this experiment, there had been concern that inversion might reduce identifiability of the stimuli as faces -- recall the adults who perceived them initially as just circles containing lines and curves. This, together with the reports by other investigators that young children recognise inverted stimuli less readily than upright ones in general, gave some grounds for anticipating negative or inconclusive results from Experiment 5. That inversion did not after all give general difficulty has been demonstrated: the percentages of errors were not appreciably higher in Experiment 5 than in Experiment 4.

A few of the Experiment 5 subjects were observed to tilt their heads sideways when viewing the stimuli or to attempt to turn round the cards (which the Experimenter prevented), suggesting these children's awareness that the forms had a "normal" orientation from which they currently deviated. There was also some evidence that the stimuli were treated specifically as <u>faces</u> even when upside-down, from comments made by some of the subjects about them during pair-comparison. The commonest verbal signs of recognition of "faceness" were: use of the word "face" itself (or sometimes, "man"); naming an attribute specifically as "mouth", "eyes", "nose"; or descriptions imputing characteristics peculiar to faces or people, such as "smiling", "sleeping". (Remember that the Experimenter had been careful to avoid using such terms herself.) With the upright stimulus presentation of Experiment 4, 27.5% of the 40 subjects gave some verbal comment indicative of faceness. Of the 27 subjects receiving inverted presentation in Experiment 5, 29.6% used language appropriate to faces. These remarkably close proportions of subjects producing face-related utterances in the two orientation conditions make it seem likely that inversion caused little interference with interpretation of the figures as faces.

The consistent dominance of Mouth in correct judgments, moreover, argues that some property peculiar to this attribute continued to be important -- that inversion did not hinder its identification to any major extent. One might then speculate that whatever it was about Mouth that gained it favour over the other attributes, it had to do with the face context within which it occurred. This possibility will be directly addressed in Parts 4 and 5. First, however, a further experiment was conducted under this section.

EXPERIMENT 6

Pair-comparison preceded by familiarisation with the stimuli

It has been argued that, by the device of equating physical structure (as far as feasible) across the elements composing the present schematic faces, their three attributes should be about equally objectively salient and their distinguishing levels equally discernible. On this basis, in other words, there is no reason to suppose any of the attribute distinctions to be more, or less, detectable than another. So correct pair-comparison judgments (or, conversely, errors) should in principle be roughly evenly distributed among the attributes.

As Experiments 4 and 5 have demonstrated, any effects of unequal visual saliencies or mirror-image reversal confusions were indeed minimal. Yet, consistently, Mouth variants were correctly discerned significantly more often than were Eyes or Nose. It begins to look as though the observed patterns of correct judgments and confusions occurred in these pair-comparison tasks not so much through a failure to detect a particular attribute or the distinction between its variants as because of a deliberate assignment of priority to one attribute over another.

The following experiment provided a further check on the detection versus subjective weighting question. As it was meant primarily as just a supplementary test, and because of the length of the experimental session, only a small sample of subjects was used -- if anything unexpected emerged, a larger sample could be taken later. The critical measure was again to be the distribution over Mouth, Eyes and Nose of correct "not the same" judgments in conventional pair-comparison with the second set of schematic faces -- exactly as in Experiment 4. But besides having all three face attributes constructed along physically identical lines, an additional step was taken to help ensure that the children could detect all the attributes and their distinguishing levels. This was accomplished by providing a kind of pretraining, intended to familiarise the children with the discriminanda they would subsequently meet in the pair-comparison task. In this preliminary familiarisation period, the child was to describe each stimulus face. Special emphasis was placed on eliciting indications that he perceived the dissimilarities among faces, particularly the distinction between the two curved variants within each attribute.

By this means, it was hoped, any attribute or distinction that might not be initially visually salient could be rendered more so, such that in subsequent pair-comparison the child would be more likely to detect it than he would have been without the prior experience. In other words, the familiarisation period was expected to have, if anything, an all-round facilitatory effect on discrimination of the attribute variants during later pair-comparison. In fact, since there was more room for improvement on Eyes and Nose discriminations, any enhancement effect should be greater for those two attributes than for Mouth, which is already near its ceiling. The distribution of correct "not the same" judgments should thus tend to level out over the three attributes -- at least according to the "differential detection" view of children's confusion errors in sameness-judgment tasks.

Subjects

Fourteen children served as subjects, eight from a nursery class in St. Andrews and six from a playgroup in Oxford. They were assigned to two groups, of four boys and three girls each, balancing age as far as possible. This subdivision was relevant only for the pair-comparison part of the test session, in which, as before, one group was to receive the stimuli in Left-Right alignment and the other in Up-Down alignment. The age range of the children in the L-R Group was 3;9 to 5;0 years and that of the U-D Group was 3;8 to 4;10 years, both with a mean of 4;5.

Materials and Procedure

The test stimuli were again the six schematic faces of the second set (shown in Appendix 1, Figure 3). The set of "stick figures" in various postures that had been employed for practice in Experiment 1 were re-used here for practice. (For present purposes, the original stick figures were preferred over the practice stimuli of Experiments 4 and 5 since those geometric forms were more difficult for children to describe verbally, and training on verbal description of stimuli and their features was now more important.)

The test session was divided into two parts, run consecutively. First came familiarisation with the stimuli, followed immediately by the pair-comparison task. The procedure during the pair-comparison part was the conventional method (with the faces in upright orientation) identical to that employed in Experiment 4, and requires no further comment here.

The results of the pair-comparison task will be presented shortly, but first the prior familiarisation period needs explaining. As already noted, the aim of this "training" was to try to make sure that the child perceived all stimulus features, especially the distinctions that would be relevant to correct performance in subsequent pair-comparison. The format was as follows.

The six schematic faces were laid out face up in a row on the table, in a different order for each child (obtained simply by shuffling the stimulus cards). This array was on view throughout the familiarisation period. A single card was taken from the row and placed centrally below it, in front of the child. The Experimenter asked the child about the stimulus figure, always beginning with, "What do you see in this picture?". At the start of a session, the child's first responses were typically rather vague ("It's a man", "A face"). Questioning proceeded, at first on a very general, open level (e.g., "Can you tell me a bit more about the face?"); but if the child did not come to describe the detail contained in the figure, the Experimenter prompted by asking more "leading" questions about specific parts of the face ("What about this bit?" [Experimenter pointing to a feature], "Can you tell me what this is?", "What does this bit look like?", etc.). Questioning was continued until the child had mentioned each attribute using either its conventional "adult" name or some idiosyncratic but consistent term, and had shown, either linguistically or gesturally, to the Experimenter's satisfaction that he could distinguish between the levels of each attribute. A transcription of a typical question-answer sequence follows, from a child aged 3;8 given face 3 on her first trial. (Appendix 3.2, pp. xvi-xviii, gives further examples of the children's responses.)

Experimenter:	Child:
What do you see in this picture?	A wee man.
Can you tell me some more about the picture?	He's got a face.
Can you tell me a bit more about the face?	He's got funny eyes -- they're sleeping.
What else has he got?	[no response]
What other things are in the face?	[no response]
What about this bit [E. points to Nose]?	A nose.
That's right. What's the nose look like?	A straight thin nose.
What else is in the face?	[no response]
What about this bit [E. points to Mouth]?	A straight mouth.

The Experimenter did not normally mention the word "face" nor name an attribute until the child had done so, after which the Experimenter adopted whatever terms the child employed. In attempting to elicit a description of a specific attribute, the Experimenter simply pointed to it and referred to it as "this bit" (etc.). The order in which the Experimenter introduced the attributes, when this was necessary, differed across subjects. For example, Eyes would be introduced first for one child, second for another, and last for another child, lest the order in which the attributes were brought to the child's notice should affect the order in which he scanned or ranked them in subsequent pair-comparison.

No child was unable to name an attribute himself, but there were certain exceptions to the Experimenter's rule of not describing an attribute variant before the child had. These were the cases where the child simply did not reply to the Experimenter's question. Most of these instances had the character of the child's being at a loss to say what he wanted -- he seemed to be "searching for the right words", unable to grasp some distinction linguistically rather than perceptually. This was particularly true of the Nose. The child could usually discriminate a "straight line nose" from a "curly nose", but might fail to describe the two curved variants in a way that distinguished them. The subjects here were too young to be able to cope reliably with the terms "left" and "right" for specifying the direction of the curve. The Experimenter therefore suggested, say, that the child point his own nose the same way as the stimulus face, or that he say something like "the nose is turned to the door" (to the child's left) versus "it's turned to the window" (to his right). This was readily accomplished by those subjects who had initial difficulty. This kind of problem was rarely encountered with Eyes and Mouth, where the children had available terms for describing the different curves (e.g., "up" vs "down").

When the Experimenter was satisfied with the child's description of a face, the stimulus card was replaced in the upper row and the next one was brought down, until all six stimuli had been dealt with. As the session progressed, the children appeared to become aware of the kind of answer expected and to offer adequately detailed descriptions with less intervention from the Experimenter.

The session using the test faces was preceded by practice on the stick figures, with which three familiarisation trials and then three pair-comparison trials were presented successively, using the test procedure. The stick figures proved easy for the children to cope with, since the distinctions among stimuli could normally be captured in terms of an action ("It's a running man", "He's kicking up his leg"). The Experimenter pressed for additional information ("Is he running towards the door or towards the window?", "What's he doing with his hands?") to accustom the children to attending to detail and to the lengthy questioning in general. All subjects performed satisfactorily in practice on both familiarisation and pair-comparison. Following practice, the schematic faces were shuffled and laid out in a row for the familiarisation part of the test session. After those six trials, the stimulus cards were reshuffled and turned face-down, ready to go immediately into the pair-comparison test trials. The entire session lasted for 30 to 40 minutes.

Prediction

The preliminary familiarisation period was not important here (though it will be discussed in Part 4). What was of interest was any effects the familiarisation with the stimuli might have on the children's subsequent performance in the pair-comparison task, which provided the test of the detection versus subjective weighting question.

The null hypothesis was that there would be no difference in frequency of correct "not the same" judgments across the attributes Mouth, Eyes and Nose, i.e., across the three critical trials, on face pairs 12, 34 and 56, respectively. If the prior experience of the discriminanda afforded by the familiarisation period had any effect it was expected to be reflected in a general improvement in detection of differences between stimuli: the distribution of correct responses would tend to level out over the attributes, a result that would be in keeping with the null hypothesis.

However, if in Experiments 4 and 5 -- without the prior experience -- the children were already noticing the differences, then the familiarisation now provided could not be of much help. The results would then be expected to parallel those of Experiments 4 and 5, i.e., Mouth would predominate in correct judgments. This was the alternative hypothesis under consideration. On the basis of those previous experiments, in fact, it was predicted that the differential frequency of correct responses among the attributes would be maintained. Such an outcome would be in keeping with the view that, in deciding about the sameness of these faces, the children were weighting Mouth over Eyes or Nose.

RESULTS

Identical pairs: Wrong "not the same" judgments. Among the 14 subjects, each receiving six identical pairs, only one false negative response occurred -- made by a five-year-old girl on pair 22 presented in L-R alignment. It was probably simply an "accidental slip".

Two-attribute distinctions: Wrong "same" judgments. Each subject received three trials in which the stimuli differed by two attributes. Again, only one error was made, this time by a four-year-old boy in the U-D Group, on pair 15 (differing in Mouth and Eyes).

Entirely correct sessions. Six of the 14 subjects answered correctly on all 12 pair-comparison trials -- three children in each stimulus alignment condition. Their ages ranged from 4;2 to 4;10 years, and both sexes were represented.

The above results suggest that the children who received prior experience with the stimuli were generally more successful than their "inexperienced" predecessors. In Experiments 4 and 5, where the stimuli were unfamiliar, errors on identical-match pairs and on pairs differing by two attributes had been quite infrequent; but now they were all but nonexistent. Moreover, the percentage of subjects managing correct responses throughout the task was considerably higher than before. Six out of 40 children in Experiment 4, and three out of 27 in Experiment 5, were entirely correct -- proportions of 15% and 11% respectively; the figure in the present case was 43%. It appears, then, that familiarising the subjects with the discriminanda did improve their chances of success when they were subsequently presented with the same figures for pair-comparison.

Critical trials: Correct "not the same" judgments. Of the relatively few errors made, almost all involved confusions of just one attribute. The crucial question is, how were they distributed among the attributes, compared with previously. Table 27 shows the percentages of children giving correct "not the same" answers on the three critical pairs, 12 (M different), 34 (E different) and 56 (N different). (For ease of comparison, the results from the preceding two experiments with unfamiliar stimuli have been carried forward. Only the percentages for both L-R and U-D conditions combined are reproduced here; for details of the subgroups separately, see Table 24 on p. 140, and Figure 5, p. 160.)

As before, subjects given the stimuli in U-D alignment performed slightly worse than those given L-R alignment. But also as before, in both stimulus presentation conditions Mouth was clearly far superior to Eyes and Nose in terms of the number of subjects who managed to distinguish the two curved variants of each attribute.

Cochran *Q* tests were applied to these data, with significant outcomes in all cases. Taking all 14 subjects together, the differential response frequencies over the three attributes were associated with a p-value of < 0.0005 ($Q = 21.273$). For the seven children in the L-R Group alone, again $p < 0.0005$ ($Q = 18.80$); but results for the seven in the U-D Group only just reached significance, at $p < 0.05$ ($Q = 4.667$). (All

TABLE 27 (Experiment 6)

Percentage of judgments that were correct ("not the same") on critical pairs differing by a single attribute, when subjects had been familiarised with stimuli. Overall results from Experiments 4 and 5 (unfamiliar stimuli) are added.

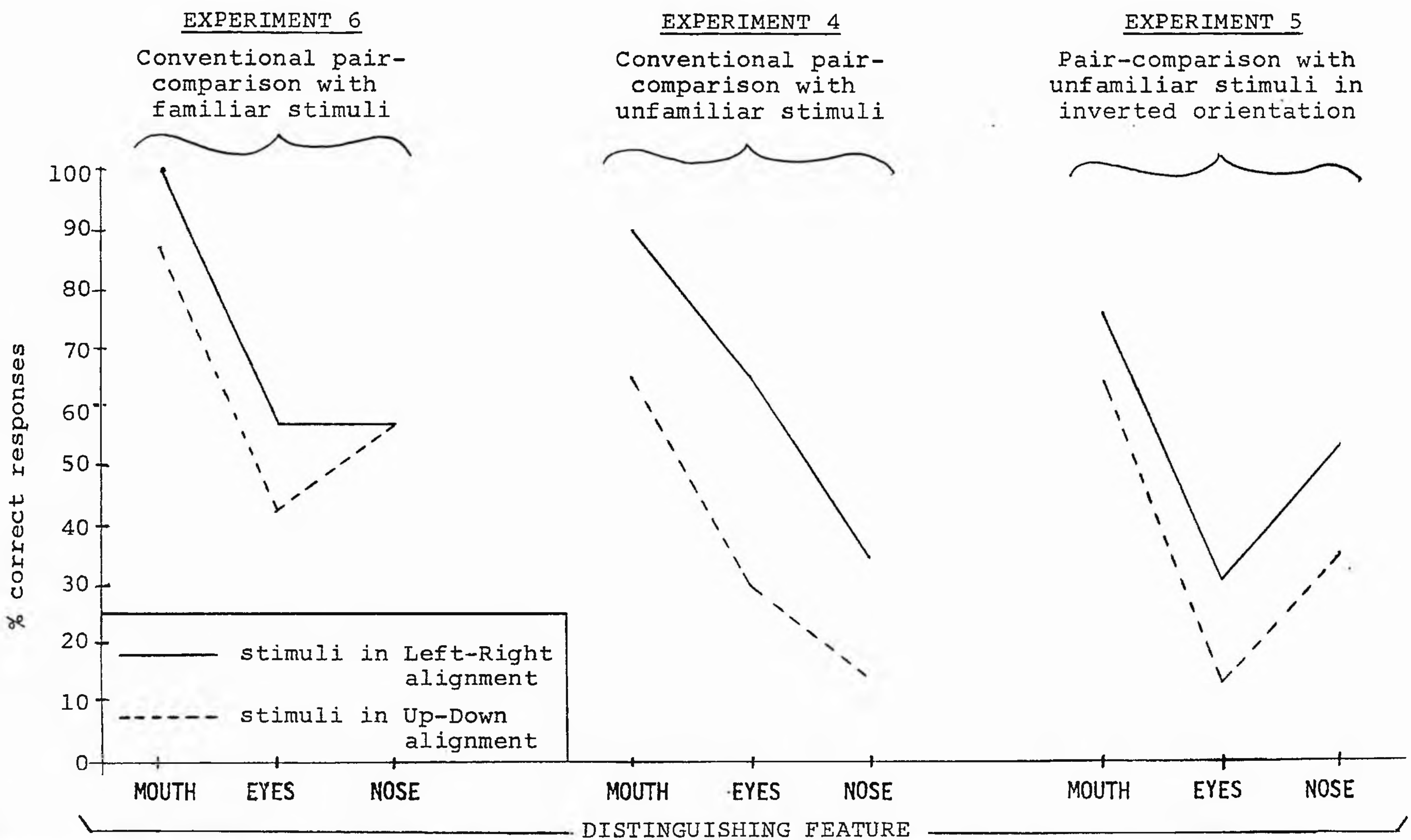
		STIMULUS PAIR AND DISTINGUISHING FEATURE		
		12 M	34 E	56 N
Experiment 6 (familiar stimuli)	L-R Group (n = 7)	100	57.1	57.1
	U-D Group (n = 7)	87.5	42.5	57.1
	All subjects (N = 14)	92.9	50.0	57.1
Experiment 4 (unfamiliar stimuli)	All subjects (N = 40)	77.5	47.5	25.0
Experiment 5 (unfamiliar stimuli, inverted)	All subjects (N = 27)	70.4	22.2	44.4

significance levels were one-tailed, since an uneven response distribution was predicted with the direction of bias specified in advance; df = 2.) Binomial tests (one-tailed) were again conducted between pairs of attributes: Mouth vs Eyes, $p = 0.016$ ($N = 6$, $x = 0$); Mouth vs Nose, $p = 0.031$ ($N = 5$, $x = 0$). These values pertain to the data for all subjects combined -- the numbers in the two subgroups were too small to allow separate treatment by stimulus alignment condition. By inspection, there was no difference between Eyes and Nose in frequencies of correct responses. The statistical analyses confirmed the marked superiority of Mouth over Eyes and Nose in correct pair-comparison discriminations.

At the outset of Experiment 6, we noted the possibility of a general improvement in pair-comparison performance arising from the children's prior experience with the stimuli during the familiarisation period. We have already seen evidence of such an improvement: in the lower proportion of errors on identical-match pairs and two-attribute differences, and in the higher proportion of entirely correct sessions achieved by the experienced subjects in Experiment 6 compared with their inexperienced counterparts in Experiments 4 and 5. Table 27 shows that a similar pattern obtained for judgments of pairs differing by a single attribute. The Experiment 6 subjects performed best and the Experiment 5 subjects worst in terms of overall proportions of correct discriminations achieved (i.e., disregarding the differential frequencies across the three attributes). This trend is highlighted in Figure 5, a graphical representation of the data from Tables 24 and 27, for S-R and U-D conditions separately. What emerges is a steady decline in the general level of performance across Experiments 6, 4 and 5. The three sets of experimental conditions -- from familiarity with the stimuli, to unfamiliarity, to unfamiliarity plus stimulus inversion -- appear as if on a

FIGURE 5

Performance decrement over various conditions of pair-comparison: Percentages of responses that were correct ("not the same") on pairs differing by a single attribute.



linearly-sloping gradient according to the degree of facilitation/ hindrance each afforded.

There is an interesting point to be noted in this connection. Occasionally on a pair-comparison trial, a child would give a quick answer but promptly change it. With respect to the trials involving a single-attribute distinction, the impression from the subjects' verbal comments was usually that the child had answered "same" rather hastily after a superficial glance at the pair, but then realised there was some difference. In the data presented so far from all experiments, when a child changed his mind his first answer was ignored; only the final judgment was scored, since it seemed more representative of the child's true decision and of his capacity for judging sameness-difference.

When the stimuli were unfamiliar to the child, reversed decisions occurred quite rarely. In Experiment 4, only two children (out of 40) made a switch, both from "same" to "not the same" on the pair differing in Eyes. In Experiment 5, three subjects (out of 27) changed their minds -- one from same to different when Eyes were distinct; one from same to different when Nose was distinct; and one from different to same when Nose was distinct. When the stimuli were familiar, however, there was a higher incidence of reversal. The 14 subjects in Experiment 6 produced six instances: two each on Mouth, Eyes and Nose, all in the direction of false "same" to correct "not the same" judgment.

Table 28 presents the data from the three experiments comparing the percentages obtained by scoring the children's <u>first</u> responses against those yielded by their final responses (the latter reproduced from Table 27). Though for Experiments 4 and 5 there is not much change whether first or last decision was scored, the picture for Experiment 6 is of a

TABLE 28

Percentage of judgments that were correct ("not the same") on critical **pairs** differing by a single attribute, comparing first against second responses (where given) across Experiments 6,4 and 5

	STIMULUS PAIR (AND DISTINGUISHING FEATURE)		
	12 (M)	34 (E)	56 (N)
Experiment 6 (familiar stimuli)	78.6 (92.9)	35.7 (50.0)	42.9 (57.1)
Experiment 4 (unfamiliar)	77.5 (77.5)	42.5 (47.5)	25.0 (25.0)
Experiment 5 (unfamiliar, inverted)	70.4 (70.4)	18.5 (22.2)	44.4 (44.4)

Notes to Table 28. Under each stimulus pair, the left-hand column gives figures based on first responses and the right-hand column (figures in parentheses) takes into account second responses, after subjects had changed their minds (cf. Table 27).

Numbers of subjects were: Experiment 6 - 14; Experiment 4 - 40; Experiment 5 - 27 (with both stimulus alignment conditions combined).

level of initial achievement markedly lower than that finally attained after minds were changed. Unfortunately, the small number of subjects participating in Experiment 6 means that minor shifts in response frequencies appear exaggerated when converted to percentage terms, relative to the shifts on the other two experiments with their larger samples. Still, the fact that 14 subjects in Experiment 6 produced six reversed decisions, while 40 in Experiment 4 produced only two reversals, is noteworthy. For, when first responses alone are considered, the children in Experiment 6 appear to have fared no better than their inexperienced predecessors. While not too much weight can be placed on this apparent result, it is suggestive that what the prior familiarisation period accomplished had to do not so much with improving discrimination directly, as with training the children to take a second look and give a more carefully considered opinion.

Although this could certainly be regarded as an effect on the children's detection of the attributes and their distinguishing variants, it is not quite of the character initially envisaged. It has not altered the differential proportions of correct judgments accruing to the three attributes by bringing performance on Eyes and Nose distinctions up to the level of performance on Mouth -- the effect looked for. Discrimination has been instead enhanced about equally for all three attributes, with the result that Mouth was still by far the most often successfully discerned. Although comparatively few subjects were tested in Experiment 6, this pattern was so pronounced, statistically robust, and so uniform across the subjects, that it was considered unnecessary after all to enlarge the sample.

Summary and conclusions

Throughout the three experiments in this section of the study, the same stimulus set of six schematic faces has been employed, in the same basic pair-comparison task format. Despite the fact that the physical structure of the variants was more or less equated across the three attributes of these faces, Mouth consistently predominated over Eyes and Nose in correct "not the same" judgments. This was true whether the stimuli were presented in conventional orientation (Experiment 4) or in inverted orientation (Experiment 5), and also when the pair-comparison trials were preceded by a kind of training period designed to familiarise the subjects with the stimuli and promote detection of any attribute that might otherwise, for some reason, be noticed less than the others (Experiment 6).

What did tend to vary across the three experiments was the children's general level of pair-comparison performance; it appears that stimulus inversion (like Up-Down alignment) slightly impaired it while provision of prior experience with the stimuli slightly enhanced it in relation to the base level of Experiment 4. The three conditions also apparently had some effect on which attribute would be the most often confused, Eyes or Nose. With conventional upright presentation, Nose was inferior, while Eyes were inferior when the stimuli were viewed upside-down. This cross-over can perhaps be explained in terms of a visual factor: the part of the stimulus figure to which gaze is first directed, and the subsequent direction of scanning from that starting point (Braine, 1965). (The improvement in judgments of Nose in Experiment 6 compared with Experiment 4 is possibly attributable to the prior training on detection and discrimination given in the former case.)

But the consistent dominance, under all conditions, of Mouth cannot be accounted for in purely visual terms. It was not that Mouth was more visually salient -- if anything, Eyes should have been the most objectively noticeable attribute, since it was composed of <u>two</u> lines. It was not that mirror-image reversal effects played any significant part -- for otherwise Eyes and Mouth scores should have been equal, and there should have been a cross-over effect between the Left-Right and Up-Down stimulus alignment conditions with regard to scores on Nose versus Eyes/Mouth. It was not that the children attended primarily to the bottom of the stimulus card, for then Eyes should have been correctly judged in the stimulus-inversion condition, when in fact Eyes were the most often confused attribute. Finally, the attempt to see whether Eyes and Nose scores could be brought up to the level of Mouth by emphasising the discriminanda during a pretraining period ended in a negative outcome -- performance on Mouth was simply equally raised.

This section of the study was embarked on with the objective of determining whether the response patterns observed in the first three experiments of the programme could be at least partly due to the fact that in the earlier set of stimulus faces the three attributes (then Shape, Eyes and Mouth) were quite different in their physical structure. Since the Shape variable was subsequently dropped, the question became whether the earlier finding of Mouth dominance could be attributed to differential objective visual salience between Mouth and Eyes. On the basis of the results of the second group of three experiments, the answer is that this is unlikely. And if the dominance of Mouth cannot be explained on such perceptual grounds as have been examined here, we are drawn towards the conclusion that a cognitive factor was involved. For some reason, Mouth was <u>subjectively</u> salient; as a criterion of sameness-difference, it appears to have been assigned more weight than were Eyes or Nose.

Why? Is the mouth generally a highly significant part of the face for young children, or is it rather that the specific nature of the present stimulus faces made it so just in these circumstances? Hints have already appeared that in their incidental verbal comments during test sessions, the children treated the attribute Mouth in different terms from the way they referred to the other features. In the following section, the children's spontaneous remarks and requested justifications for their decisions in the preceding experiments are examined in more detail, with a view to attempting to clarify this issue, among others.

PART 4

DETECTION OR SELECTION OF STIMULUS ATTRIBUTES?

II. SOME CLUES FROM SUBJECTS' JUSTIFICATORY RESPONSES

During Experiments 2 to 6, alongside the subjects' main task decision about the sameness of stimuli on each trial, the Experimenter recorded any instances where the child commented on a stimulus figure or indicated specific aspects of sameness or difference between figures. Fairly often, such information was volunteered spontaneously. Additionally, in Experiments 2 and 3 subjects were explicitly asked at random intervals to comment or to "justify" a decision ("How are they the same/not the same?", "Can you see anything else about them that's not the same?", etc.).* For the sake of simplicity, all these responses -- spontaneous or requested, verbal or gestural -- are here called <u>justificatory</u>, as distinct from the <u>main-task</u> responses (pair chosen from triad, or pair-comparison judgment) with which they were associated.

The justificatory responses were all of rather similar nature: indicating that one or more attributes shared (or

---

* Experiment 2 was embarked upon with the intention of eliciting a justification on every trial. This proved infeasible, partly because the test session became too lengthy and partly because some of these young children were reluctant or even became upset, reacting as though they felt the correctness of their main-task decision was in question. In any case, justifications were quite consistent within a subject's session, so that a sample from each child sufficed. In Experiment 3, therefore, justifications were requested on only a few trials at random (with reluctant children not questioned after an initial attempt), and thereafter were not specifically asked for at all.

did not share) the same level in the given stimulus pair by pointing to the feature, naming it, or describing its (visual or other) character. They were thus relatively easy to analyse. Where they did vary across children and across stimulus pairs, it was largely in terms of:

the amount of information included in the response (e.g., number of attributes noted);

order in which attributes were referred to in multiple responses, i.e., when more than one was noted within a single trial;

number of times each attribute was referenced;

and manner of referring to an attribute (e.g., by just pointing, or by verbally describing its variants, etc.).

These quantitative and qualitative factors will be considered in turn for their relationships, if any, with the patterns of correct matches and confusion errors observed in main-task performance.

The aim of examining the justificatory responses was to try to shed further light on such questions as the extent to which, and respects in which, young children succeed or fail in judging sameness/difference, how they go about making their decisions, and ultimately what they understand by the word "same" in these tasks. In the present context, the important finding to be explained is that children's confusion errors are not randomly distributed, but tend to be systematically biassed towards certain types of discriminanda. Here, two main possibilities are under consideration: that particular stimulus features were overlooked either because they had low objective visual saliency relative to other features and were therefore accidentally not noticed (failure-to-detect view); or because they were treated as not relevant in the given task and stimulus context, were assigned a lower subjective weight than other features, and were therefore deliberately ignored (failure-to-select view).

On the evidence so far from the present series of experiments, it has been suggested that the observed confusion errors seem to owe more to non-selection than to non-detection. Is there anything in the justificatory responses to help support or refute that contention? And if the former, are there any clues as to the possible bases on which differential subjective weights were assigned among the stimulus attributes? To illustrate the idea: Suppose a Shape-matching subject explicitly and consistently indicates identity of Shape as the reason for his decisions and makes little or no reference to Mouth. In that case, we have no way of telling whether or not he perceives the Mouth differences that he confuses. However, if he does explicitly (and correctly) point out sameness/difference of Mouth, then we are forced to conclude that Mouth is perceived but not used as the basis for decision. Further, given the Shape-matcher who indicates that he perceives both Shape and Mouth, we can ask whether there is any difference in the way he refers to them (e.g., in the terminology of his verbal descriptions) that could tell us what it is that has led him to base his sameness judgment on Shape alone.

The various methods of analysing the justificatory responses were largely parallel across all experiments. However, Experiments 2 and 3 will be discussed separately from Experiments 4 - 6 because of the two distinct stimulus sets employed, the attributes in the first set being Mouth, Eyes and Shape (M, E, S) and in the second, Mouth, Eyes and Nose (M, E, N) (with Eyes constructed differently in the two cases).

It should be noted that the aim of recording justifications was not to furnish independent evidence that would stand on its own. While they may (or may not!) provide helpful information, they were intended simply as an adjunct to the principal data obtained from the corresponding triad and pair-comparison judgments. Their treatment in what follows is therefore informal.

Samples of the subjects' justificatory responses are reproduced in Appendix 3.

1. Experiments 2 (triads) and 3 (pair-comparison)

See Figure 1 of Appendix 1 for the composition of the eight schematic faces making up the stimuli employed in these two experiments.

No discernible differences in the justifications emerged according to whether the subjects received Framed or Unframed stimuli; stimulus condition will therefore be ignored throughout the following analyses. Instead, the subjects were divided into their classes according to the dominant attribute matched in the main task. The triads task yielded 12 Shape-matchers, 11 Mouth-matchers, one Eyes-Matcher, and 12 Mixed-matchers (unclassified because inconsistent). (There was no possibility of being correct in this task, recall.) All of these subjects produced at least several justificatory responses. Of the 62 subjects taking part in the pair-comparison experiment, however, 22 produced no justificatory responses; these were excluded from the present analyses, except in the calculation of basic response rates. The remaining 42 subjects were distributed among matcher-groups as follows: 4 Shape-, 9 Mouth-, 4 Eyes-, 4 S+M- (i.e., for whom both Shape and Mouth had to match, though not Eyes, before stimuli were judged to be the same), 2 S+E-, 2 E+M-, 6 Mixed-, and 9 Correct-matchers. (For present purposes, subjects who made fewer than four confusion errors altogether were classed as if entirely Correct. See Table 19, p. 92 , for the original distribution of the entire sample of 62 subjects.)

Samples of the kinds of justificatory responses produced in Experiments 2 and 3 are provided in Appendix 3, Section 1. Pages ii - vii show sequences from the response sheets covering a series of trials.* These segments were selected to provide a general picture of the sorts of

---

* For ease of comparison, these segments have been transcribed such that the order of the triads is the same across subjects. They do not represent consecutive responses as actually given, since order of trials (and position of stimulus cards within trials) was randomised across subjects.

responses given. They are fairly typical of the range of verbal responses that did occur, but are not necessarily typical of the record sheets themselves since, as we shall see, many responses consisted of simply pointing to or monosyllabically naming an attribute and since -- particularly in pair-comparison -- many trials received no justification at all. Pages viii - xi of Appendix 3, Section 1 list some further individual justificatory responses, again often atypical in their complexity, or added specifically to illustrate some peculiarity.

Response rates*

Overall, some justificatory response was supplied on half of the trials in the triads, but only a quarter of the time in the pair-comparison task. This difference in rate of responding was presumably due simply to the restriction of requests for justifications to a limited sample of trials in Experiment 3.

Response frequencies varied somewhat across the different classes of subjects. Table 29 shows the proportions of trials with which were associated no response, reference to just one attribute, and reference to two or to all three attributes, for each matcher-class. In both triads and pair-comparison tasks, subjects who performed without any consistently dominant attribute (Mixed groups) offered

---

* For all experiments, responses consisting simply of, e.g., "They're the same/not the same", where the only reference was to the entire stimulus figure rather than to a particular attribute, were classed as "no response" as far as justifications were concerned. Such responses merely answered the test-question of the main task without justifying those main-task decisions.

TABLE 29

Percentage of trials receiving no response, or reference to one, two, or all three attributes, by class of subject

CLASS OF SUBJECTS	(n)	NUMBER OF ATTRIBUTES REFERRED TO PER RESPONSE			
		No response	Ref. to 1 attribute	Ref. to 2 attributes	Ref. to 3 attributes
Triads task					
Shape-matchers	(12)	75	17	6	3
Mouth- "	(11)	65	23	12	0
Eyes- "	(1)	82	18	0	0
Mixed- "	(12)	19	45	28	8
Pair-comparison task					
Shape-matchers	(4)	86	8	5	2
Mouth- "	(9)	73	21	4	2
Eyes- "	(4)	88	10	1	1
S + M- "	(4)	63	23	11	3
S + E- "	(2)	82	16	2	0
E + M- "	(2)	93	7	0	0
Mixed- "	(6)	61	14	11	14
Correct- "	(9)	85	8	4	4

the greatest number of justificatory responses (seen from their low scores in the "no response" column), and made reference to more than one attribute within a trial more often than did the other subject classes. Mouth-matching subjects also produced fairly high response rates.

In Experiment 2, where the task was to select the most similar pair from a triad, the children sometimes spontaneously noted the third, unchosen face in their justificatory responses. These references were of two types:

(a) "All the same" responses. The child said that the three faces in the triad were "all the same" (and usually had to be encouraged to find a pair "more the same" than the others). (For illustration, see Appendix 3.1, pp. viii - xi, examples 6 - 9, 11, etc.) This happened on trials where the child's dominant matching attribute was common to all three faces, e.g., when a Shape-matcher was confronted with a triad in which the three possible pairs shared just Shape, or S + M, or S + E. Over the 56 trials comprising this task, an "all the same" response could occur on the basis of shared dominant attribute a maximum of eight times for each attribute. In fact, this response was given a mean of four or five times per subject, with several children attaining the maximum. This relatively high frequency suggests the "power" of the dominant matching attribute; it happened even when the three faces appeared (to an adult!) rather dissimilar (compare, e.g., faces 1, 4 and 7 -- straight Mouth common -- in Figure 1 of Appendix 1).

(b) Comparison with third face. Having selected a pair, the child commented, directly or indirectly, on one or more attributes in the third face, relating or contrasting it with the chosen ones. Appendix 3.1, pp. viii - xi gives a number of examples, e.g., 10, 18, 19, 21, and especially 14 for its detail. Comparison with the third face occurred on average only twice per 56-trial session for Shape- and Mouth-matching subjects, but seven times for Mixed-matchers. Given their higher basic rate of giving justifications and multiple responses as well, one might speculate that Mixed-matchers were inconsistent because they were more aware of all the attributes and of alternative possibilities.

Order of Reference

Multiple responses -- those containing reference to more than one attribute within a single trial -- were examined to see whether there were any consistent patterns of order of reference to the three attributes. One might, intuitively, expect subjects to refer to their particular dominant attribute first. Or is it rather that one attribute tended to receive first mention across subjects regardless of what the child actually matched?

Six orders of reference were possible: S>M>E, S>E>M, M>S>E, M>E>S, E>S>M and E>M>S. Table 30 shows which of these actually occurred most often. (For simplicity, in responses referring to just two attributes, the unmentioned one was assigned last place in the sequence.)

TABLE 30

Most common orders of reference to the three attributes in multiple responses

TRIADS TASK		PAIR-COMPARISON TASK	
CLASS OF SUBJECT	MOST COMMON ORDER OF NOTING ATTRIBUTES	CLASS OF SUBJECT	MOST COMMON ORDER OF NOTING ATTRIBUTES
Shape-matchers	M>E>S OR S>E>M	Shape-matchers	E>S>M
Mouth-matchers	M>E>S	Mouth-matchers	M>E>S
		S + M -matchers	M>S>E
Mixed-matchers	E>M>S	Mixed-matchers	M=E>S
		Correct-matchers	M=E>S

Notes:

"> " means "was indicated before".

"M=E>S" means Mouth and Eyes occurred equally often in both first and second place, while Shape was always last.

Omitted classes of subjects -- those matching by Eyes or by the S+E or E+M configurations -- produced zero or at most two multiple responses.

It is clear from the table that an internal feature was typically indicated first -- even by Shape-matching subjects! In fact, Shape consistently came last rather than second in the sequence, except for those subjects taking Shape as the (or a) dominant attribute in the main task.

Mouth-matchers were the most internally uniform group, showing almost universal preference for the order M>S>E. Surprisingly, Shape-matchers (rather than Mixed-matchers) were the least uniform -- particularly in the triads task where, for Shape-matchers, order tended to be associated with whichever attributes were actually shared by the chosen face pair. That is, on trials where Mouth differed, the common Shape was indicated first and the Mouth difference last. But when Mouth too was common, Shape was noted only rarely, and then last. These children's reactions to further questioning on Mouth-same pairs gave the impression that they took the common Shape for granted. Apparently, they chose instead to remark on the fact that Mouth, too, was shared, and seemed surprised when then asked about Shape, as if its identity was obvious.

Distribution of references among attributes

The total number of times each attribute was referred to was calculated over all responses. (Multiple mentions of one and the same attribute within a single trial were counted as one reference to that attribute; mention of two attributes counted as two references, one for each attribute.) Table 31 shows how the references were distributed among the three attributes for each class of subject. "Erroneous" references are included in these data, with their occurrence marked by a superscript _a_, _b_ or _c_, denoting the proportion of entries in the cell that were erroneous. These were instances where the child indicated (by word or gesture) that an attribute was the same in two faces when it in fact differed; or where he indicated that an attribute was different when it was actually identical (a rare occurrence); or where he gave a false verbal description of an attribute level (e.g., calling black Eyes "white"). Descriptions that were idiosyncratic but consistently applied in a manner that successfully distinguished attribute levels (e.g., systematically calling white Eyes "orange" and black Eyes "blue") were not classed as erroneous.

The right-hand column of Table 31 summarises these data, showing the order of frequency of reference among the attributes, from most to least often mentioned. The pattern rather closely resembles that of Table 30, presenting the order in which attributes were introduced in multiple responses. That is, Shape consistently received the fewest references, except by subjects who took Shape as dominant attribute in matching performance. But even for Shape-matchers internal features were prominent, receiving at least as many references as Shape. A possible explanation for this pattern will be suggested in the following subsection, when _types_ of reference are considered.

TABLE 31

Percentage of trials containing reference to each attribute, within each class of subjects

CLASS OF SUBJECT*	(n)	% references to: SHAPE	MOUTH	EYES	total refs.	ORDER OF FREQUENCY OF ATTRIBUTE REFERENCES
Triads task						
S - matchers	(12)	34	37[a]	30[c]	245[b]	M = S = E
M - matchers	(11)	19	56	25[a]	298	M > E > S
Mixed-matchers	(12)	16	45[a]	39[a]	759[a]	M > E > S
Over all subjects		20	46	34	1302	
Pair-comparison task						
S - matchers	(4)	32	32	36[b]	25[a]	E = S = M
M - matchers	(9)	17	63	20	86	M > E = S
S+M- matchers	(4)	32	63	5	59	M > S > E
Mixed-matchers	(6)	25	37	38	130	E = M > S
Correct-matchers	(9)	17	41[a]	42[a]	68	E = M > s
Over all subjects		24	48	28	368	

Key:

a. 5 - 10% of references in this cell were erroneous

b. 10 - 15% of references in this cell were erroneous

c. 25% of references in this cell were erroneous

(see text for further explanation)

">" -- received more references than

"=" -- received as many references as

* Eyes-, S+E-, and E+M-matcher groups are omitted because they produced too few references to attributes (16 or fewer total per group).

Although internal features received more references however, it was here -- especially on Eyes -- that most errors of reference were made. Shape itself was hardly ever wrongly referenced; but paradoxically, it was the Shape-matchers who produced the highest proportions of erroneous references (mostly on Eyes). Interestingly, even those children who performed entirely correctly in their sameness judgments (applicable to pair-comparison only) still made errors when indicating points of sameness/difference in justifying those judgments! Since the achievement of 100% successful discrimination over 28 nonidentical pairs can hardly be a chance outcome, these children must have both perceived and used all attribute-variants in making their judgments. It is therefore likely that their confusions of attribute levels in subsequent justifications were due to accidental "slips of the tongue" or lapses of attention. It is difficult to extend this explanation to the other subject groups, however, for one would expect proportions of erroneous references then to be about equal among these groups, whereas in fact Shape-matchers "lapsed" considerably more often than did Mouth-matchers.

Types of reference

Each reference to an attribute was classified according to the following category system (established on a post-hoc basis after inspection of the various means used to refer to attributes):

Point/Name: The child pointed to or named an attribute to indicate that it was the same (or different) across faces. He did not explicitly differentiate the two levels of the attribute. Examples of verbal responses assigned to this category are: "The mouths the same", "Not the eyes", "That bit's the same [pointing to Shape]".

Feature-level: The child verbally described the attribute in a way that explicitly differentiated between its two levels. But the description was confined to the physical, visual properties of the attribute itself. Idiosyncratic descriptions that preserved differentiation and referred to physical structure were permitted. Examples are: "That one's got black eyes and that one's got white eyes", "An upright egg and a sideways egg [= vertical/ horizontal Shape]", "They both got straight-line mouths".

Face-level: The two levels of an attribute were explicitly differentiated but in a way that went beyond physical properties, and went beyond the attribute itself, to characterise the whole face (or "the person behind the face") as being in some state or action. Examples: "They're both happy faces" (re. curved-up Mouth), "He's bulging his cheeks out like this" (child puffs out his own cheeks; re. horizontal Shape), "They do that when they're crying" (re. blank Eyes).

Further examples of the kinds of terms used in descriptions categorised as Feature- or as Face-level are provided in Table 32.

TABLE 32

Examples of justificatory responses that successfully differentiated the two levels of an attribute

(a) at the Feature-level:

SHAPE: attribute called: head, face, shape, egg

vertical orientation	horizontal orientation
upright; standing up; straight; long; thin†; round*	flat; lying down; sideways; short; fat†; not round*

EYES: attribute called: eyes

black circles	blank circles
black; blue; filled-in; coloured(-in); scribbled in	white; not coloured(-in); empty; not got scribbles in; round*

MOUTH: attribute called: mouth, face

straight line	upward curve
straight; line; straight-line; along the way; across the way; not a smile†	curve(d); curl(y); not a line; round; up the way; smile†

(b) at the Face-level:

sad; cross; grumpy; angry; frowning; crying; thin†	happy; smiling; smiley†; laughing; fat†

NOTES:

* round: This could occur in various contexts. With respect to Shape, some children said round for vertical oval, distinguishing it from (e.g.) not round for horizontal oval; or, others said sort of round for vertical and round for horizontal, etc. These instances of successful differentiation of Shape are distinct from cases where the child did not differentiate the two orientations, e.g., using sort of round for both levels. (The latter sort of reference was coded as Point/Name, not Feature-level.) Similarly with respect to Eyes: sometimes round (= blank) was contrasted with black or filled-in; but sometimes round was used indiscriminately for either level. In Mouth references, round was used exclusively for only the curved variant.

† fat/thin; smile (smiling, smiley): Whether these were coded as Face- or as Feature-level depended on their linguistic context. Thus, "He's got a fat/thin head" was classed as Feature-level (regarded as restricted to the attribute and to physical characteristics), while "He's fat/thin" was classed as Face-level (taken to refer to some state the "person" was in). Similarly, "It's got a smile", "That's a smiling mouth", etc., were taken as Feature-level, but "He's smiling" or "It's a smiley face" were treated as Face-level. (See further text pp. 83-84.)

Face-level qualifiers like sad, crying, usually referred to the straight-line Mouth but occasionally were applied to blank Eyes. The referent was normally clear from the rest of the child's utterance and from his pointing.

Other: The response was incomprehensible or conveyed insufficient information to allow classification, and the child did not clarify it. Very few responses, in fact, were unclassifiable; almost all those that did occur are listed in Appendix 3.1 (Examples 24 - 26 and 42 - 46).

Some notes on the classification procedure are in order. Firstly, cross-referencing over trials was often required before a particular response could be coded. For example, some children called Shape "round" regardless of orientation; such references were entered in the undifferentiated Point/Name category. But others said (e.g.) "round" on trials involving vertical Shape, contrasting this on other trials with "flat" for horizontal Shape. Also, some children used face to mean the whole stimulus figure, some to mean Shape, and some to mean Mouth. Thus, e.g., "They've both got straight faces" was highly ambiguous unless compared against another trial, when the child might say, "That's a straight and that's a round face" (indicating the two Shapes) or, "That's a straight face and that's a smiley face" (indicating the two Mouths).

Secondly, the distinction between Feature-level and Face-level responses could often be made on the basis of the presence of the words has (got) versus is. Thus: The face HAS (GOT) a particular FEATURE (black eyes, curly mouth, etc.) could be contrasted with The face (person) IS in some STATE/ACTION (he's smiling). Additionally, replacement of the name for the attribute by the word "face" suggested Face-level coding, provided its qualifier went beyond the attribute's physical properties. For example, "That's a happy face" was classed as Face-level because a state was ascribed to the face. But "They're both straight faces" (referring to Mouth) was classed as Feature-level despite the occurrence of "face", because the child did not explicitly transcend the attribute's physical characteristics in that particular response. Descriptions of Shape in terms of "long face", "wide face", etc., were taken as Feature-level (referring to the attribute's physical properties), but "He's fat", "He's thin", were taken as Face-level (ascribing a state to the person). Similarly, "Got a smile", "The mouth is happy" were treated as restricted to the Feature, while "He's smiling", "It's a happy face" were treated as Face-level.

Finally, each attribute was coded only once per trial, even if it received multiple references within a single response. For this purpose, the categories were treated as if "hierarchical" in the order Face-level $>$ Feature-level $>$ Point/Name $>$ Other (where "$>$" means "took precedence over"). Take, for instance, a response to face pair 15: "They've both got upright heads but one's a cross face and one's got a curl on the mouth". Upright head captured the Shape distinction at the physical Feature-level (contrasted, for this subject, with lying-down head on other trials). Curl distinguished the curved Mouth, again at the physical Feature-level. But cross face, associated with straight

Mouth, belonged to Face-level. The entire utterance was thus scored as one Shape reference at Feature-level and one Mouth reference at Face-level (no Eyes reference).

This "hierarchical" treatment of the categories does not mean to say that one type of reference was regarded as "superior to" another type. Though Face- and Feature-level were favoured over Point/Name references because they gave more information (showing that the child differentiated attribute levels), Feature-level descriptions could be regarded if anything as more "adult-like" than Face-level ones. However, we are not here interested in the sophistication of the children's descriptions for its own sake. The focus is on the association (if any) between the patterns of sameness/difference judgments made in the main task and the kinds of justifications given for those judgments. The idea is that there may be something in the way the references to attributes were linguistically encoded during justification that might tell us about the way the stimuli were cognitively encoded during the decision-making process. Face-level descriptions were allotted primacy in the classification because, as we shall see, they seemed the more interesting in this connection.

Tables 33 and 34 summarise these data, the former showing the distribution of references among the type-categories according to subject classes, the latter showing the distribution according to the particular attribute referred to, irrespective of subject class.

TABLE 33

Percentage of references that were of each type, for each class of subject

CLASS OF SUBJECT (n)	TYPE OF REFERENCE				
	Point/ Name	Feature-level	Face-level	Other	Total refs.
Triads task					
Shape-matchers (12)	61	28	9	2	247
Mouth- " (11)	62	21	17	0	298
Mixed- " (12)	56	29	13	2	761
Pair-comparison task					
Shape-matchers (4)	84	8	8	0	25
Mouth- " (9)	23	33	42	2	88
S + M- " (4)	17	63	20	0	59
Mixed- " (6)	43	38	18	0	130
Correct- " (9)	33	40	22	4	72

Notes:

Eyes-, S+E-, and E+M-matchers are omitted because these groups gave too few responses. The references they did make fell about half in Point/Name and half in Feature-level.

TABLE 34

Percentage of references to each attribute that were of each type (regardless of class of subject)

ATTRIBUTE REFERRED TO	TYPE OF REFERENCE			Total refs.
	Point/ Name	Feature-level	Face-level	
Triads:				
Shape	78	20	3	256
Mouth	54	18	28	605
Eyes	54	46	0	447
Pair-comparison				
Shape	35	65	0	95
Mouth	31	19	50	183
Eyes	47	53	0	122

Notes:

Responses by Eyes-, S+E-, and E+M-matchers are included in the above.

"Other" responses are omitted (since they could often not be linked to a specific attribute).

In the triads task, for all subject classes, more than half of the references were of the simple, undifferentiating Point/Name type. The difference among subject classes lay mainly in the relative proportions of Feature- and Face-level descriptions, with Mouth-matchers making almost as many Face- as Feature-level references, and the other two groups preferring Feature-level. As Table 34 demonstrates, however, almost all the Face-level references in fact pertained to Mouth, regardless of subject class. In general, then, Shape was referred to mostly by Point/Name, Eyes references fell about equally in the Point/Name and Feature-level categories, and Mouth received either just a Point/Name reference or a full Face-level description. Thus, in this task, type of reference was rather closely tied to kind of attribute referenced. That is, the apparent variation among subject classes seems to be largely a function of the particular attribute most often referenced, rather than of the <u>way</u> attributes were referenced.

The pair-comparison format produced greater variety among the subject classes and a stronger separation of Mouth (and Mouth-matchers) in the direction of Face-level descriptions. Shape-matchers almost always merely pointed to or named an attribute, be it Shape, Mouth or Eyes. (Again, they referred to all three attributes about equally often, though producing few responses in total.) Mouth-matchers focussed rather strongly on Face-level Mouth descriptions, and used Feature-level description rather than just Point/Name when they did refer to Shape or Eyes. Shape+Mouth-matchers, too, rarely just pointed to or named an attribute, while Mixed- and Correct-matchers used a mixture of all means of reference, except that Mixed subjects described Shape by Feature-level whereas Correct subjects merely indicated Shape by pointing.

As in the triads task, then, the Face-level descriptions pertained (exclusively, this time) to Mouth; indeed, half of the references to Mouth were made at the Face-level. Eyes references, also as before, occurred in both Point/Name and Feature-level categories. These distributions were largely independent of subject class. Surprisingly, however, Shape- and Correct-matchers (i.e., children who judged Shape correctly in pair-comparison) referred to Shape only by undifferentiated Point/Name, while the other subject groups (who confused Shape in pair-comparison) successfully differentiated its variants at the Feature-level in their justifications!

Almost all of the "erroneous" references that occurred (see pp. 178-180 and Table 31) were associated with the Point/Name type of reference. All but two of these erred in the direction of indicating that an attribute had the same level in two faces when it actually differed. Erroneous references to Mouth (which were relatively uncommon) typically consisted of the child's saying something like just "The mouth is the same", which <u>might</u> be construed as an "accidental oversight". But this kind of error never occurred on Shape, and rarely on Eyes. A number of the Eyes errors consisted of false descriptions of the physical Feature-level -- calling black Eyes "white", or *vice versa*. Again, these might be taken as accidental slips or, given the young age of some of the subjects, they might reflect difficulty in the application of the terms *black* and *white*.

However, many of the Eyes errors, and all of the Shape errors, were of another type. Here, the child used an adjective picking up on a visual aspect that was *irrelevant* in the task context (and failing to differentiate variants). The most common irrelevant description for both Shape and Eyes was "They're both *round*" (or, for Shape, also "They're both eggs"). Children who gave this kind of ambiguous description for both Shape levels tended to turn their heads sideways to view the horizontal ellipse, or attempted to turn the stimulus card round. This was surely clear indication that they *perceived* the distinction between the two Shape orientations, but regarded it as indeed irrelevant.

Summary and discussion

Table 35 presents a condensed overview of the main points dealt with concerning the justificatory responses produced in the triads and pair-comparison tasks with the first set of stimuli. Some differences among the classes of subjects and among the attributes begin to emerge.

The sharpest separation is between Shape-matchers and Mouth-matchers, in terms of subject class; and also between

TABLE 35

Summary of patterns in justificatory responses, Experiments 2 (triads) and 3 (pair-comparison)

(a) by subject classes

SHAPE-MATCHERS	MOUTH-MATCHERS	MIXED-MATCHERS
offer fewest justifications	offer fair amount of justifications	offer many justifications
don't reference dominant attribute (Shape) first in sequence in multiple responses	reference Mouth first, Shape last, in multiple response sequences	reference Mouth or Eyes first, Shape last, in multiple response sequences
reference all 3 attributes about equally after	reference Mouth most, Shape least	reference Mouth or Eyes most, Shape least
make most erroneous references	most few erroneous references	make some erroneous references
refer mostly by Point/Name without differentiating level	refer to Mouth at Face-level, and to Shape and Eyes at Feature-level	refer to Shape and Eyes by Point/Name or Feature-level; to Mouth by Point/Name or Face-level
are least internally coherent as a group	are highly internally coherent as a group	are moderately internally coherent as a group

(b) by attributes

SHAPE	MOUTH	EYES
gets fewest references	gets most references	gets fair amount of references
usually comes last in multiple response sequences (except for Shape-matchers)	often first in multiple response sequences, rarely last	comes second or first in multiple response sequences
gets few erroneous references	gets some erroneous references	gets many erroneous references
referenced by Point/Name (S-matchers) or Feature-level (other subject-classes)	referenced by Point/Name or Face-level (independent of subject class)	referenced by Point/Name or Feature-level (independent of subject class)

Shape and Mouth, in terms of attribute. Shape-matchers emerge as either less willing or less able than the other groups to justify their sameness judgments, giving fewest references, giving fewest verbal descriptions differentiating attribute levels, making most erroneous references, and showing no particular preference for their dominant matching attribute over Mouth or Eyes in terms of both frequency and primacy of mention. Mouth-matchers' justifications focussed rather strongly on their dominant attribute, although they showed that they were capable of correctly differentiating Shape and Eyes too. Mixed-matchers were the most responsive, giving many and informative justifications and including all three attributes in their references.

These patterns in the justificatory responses are consistent with the main task performance of the three subject classes. Shape-matchers, recall, tended to be not quite fully "strict" (as defined on pp. 73 - 74 and 90 - 91) in their adherence to their dominant attribute in both triads and pair-comparison performance, in that their matches were occasionally based on Mouth instead of or as well as Shape. A greater proportion of the Mouth-matchers adhered entirely strictly to their dominant attribute. And a number of the Mixed-matchers lacked a dominant-attribute matching pattern because they alternated between judging on the basis of <u>number</u> of shared attributes and judging on the basis of <u>type</u> of shared attribute, sometimes favouring a match of <u>any two</u> attributes over a match based on just one, dominant attribute.

How do the justification patterns relate to the characteristics of the attributes themselves? As Table 35 shows, Shape was associated with the smallest number of references and did not come first in the sequence when more than one attribute was referred to. Mouth, by contrast, consistently received the highest proportion of references, and was often first mentioned in multiple responses, with Eyes next.

The amount of attention devoted to Eyes in the justificatory responses contrasts with their infrequent appearance as the basis of judgment in performance on the main task. To receive such a number of references, Eyes, as a <u>general</u> feature, must have been reasonably salient and not simply totally overlooked. Remember, however, that a fair proportion of the references to Eyes in the justifications was erroneous -- they were indicated to be the same across two faces when they actually differed. It is possible, then, that the distinction between the two <u>levels</u> of Eyes was not so readily discriminable.

The case for Shape is slightly different. As we have seen, children who did refer to Shape at all tended to do so by simply pointing to or naming it as (not) the same without explicitly, linguistically differentiating between its two levels. There is a likely reason for the relative lack of Shape references and its loading in the Point/Name category that is independent from the detection versus selection issue, for Shape seemed to pose a vocabulary problem. In many instances, a child would gesture vaguely at the stimuli while appearing to be searching for words to convey what he wished to say. If the Experimenter then supplied a term like "shape" or "head", or an adjective "long" or "fat" to describe the two levels, the child's reaction implied "That's what I was trying to say!" and he would proceed to use these terms appropriately himself. It seems plausible, then, that Shape received comparatively few references simply because some of these young children lacked the lexical repertory necessary to name the attribute and describe verbally its two orientation levels. In any case, we have here no clear evidence that the children did not perceive the Shape variation.

There was no such problem in referring to Mouth, the most frequently-mentioned attribute for the majority of children. The distinction between its curved and straight-line variants could be readily captured in language in the opposition between "smiling" or "happy" on the one hand and "sad" or "straight" on the other. The striking thing here was the high proportion of Face-level descriptions: the adjective was frequently extended beyond the attribute to the whole face or the "person". In practice, the tendency was for the curved variant to elicit mostly Face-level descriptions, while the straight variant sometimes instead elicited physical Feature-level description.

However, the word "face" did often occur in connection with the straight Mouth variant -- but in a peculiar, "non-adult" context. It is one thing to say "It's a Happy face"; it is another to say "It's got a straight (<u>or</u>, flat) face" when what is meant is that the face has got a straight (flat) mouth. Yet this kind of statement occurred with surprising frequency. (Examples appear in Appendix 3.1: p. vi, Eyes-matcher on pair 18; p. vii, Correct-matcher on pairs 16 and 23; p. viii, no 2; p. x, no. 37.) Similarly, responses such as "They've got the same faces" accompanied by a point to just the Mouth were common. The children thus often treated "mouth and "face" as synonymous in contexts where an adult would not do so.*

In short, the distinction between the two levels of Mouth was obviously readily discriminable, both perceptually and linguistically, and to subjects for whom it was not the dominant matching attribute as well as for Mouth-matchers. Clearly, moreover, the structure of the Mouth variants especially lent itself to Face-level treatment. And when Mouth level was shared in a face pair, the pictures could be judged to be the same "because they're both happy/sad faces". Neither Shape for Eyes could be used to characterise identity or difference between <u>entire</u> faces in this succinct, global fashion. Perhaps Mouth may have been especially subjectively (cognitively) salient in these stimuli because sameness and difference could be so simply encoded by

* This usage of the word "face" in connection with straight Mouth is quite distinct from its usage to refer to Shape -- as in "round face" vs "straight face" for horizontal vs vertical ellipse -- which is acceptable "adult talk". Whether applied to Shape or to Mouth, references of the "straight face" kind were coded as Feature-, not Face-, level, since they adhered to <u>physical</u> properties.

reference to it? Let us then proceed to examine what happened when physical structure, and hence objective visual salience, was equated across attributes.

2. Experiments 4, 5 and 6 (pair-comparison with stimuli conventional, inverted, or familiar)

For the stimuli used in this section of the study, see Figure 3 of Appendix 1, and pp. 116 -119 . To recap: The pair-comparison method was employed throughout; stimuli were presented in conventional orientation in Experiments 4 and 6, but upside-down in Experiment 5. In Experiment 6, pair-comparison was preceded by a familiarisation period in which the subjects were required to describe each stimulus figure verbally; these descriptions will also be discussed here. The presentation of stimulus pairs under the two alignment conditions, Left-Right and Up-Down, will be ignored since they were associated with no obvious differences in justifications.

The three stimulus attributes were now Mouth, Eyes, Nose (M, E, N), with the outer head-shape a constant circle. Each attribute had three levels: one a straight line (neutral variant) and two, mirror-image, curved variants. The three critical stimulus pairs were 12 (differing in just orientation of M curve), 34 (differing in just E curve) and 56 (differing in just N curve). All other pairs either differed by two attributes or were identical.

The patterns of correct "not the same" judgments were roughly parallel across the three experiments, with Mouth receiving by far the greatest number of correct discriminations -- more than half as many again as Eyes or Nose (see Table 27, p. 158 , and Figure 5, p. 160) for further detail). The main question here is then whether there is some corresponding pattern in the justificatory responses that might bear on the main task findings that the two curved Mouth variants tended to be successfully discriminated while the two similarly-constructed curved variants of Eyes and Nose were often confused.

Section 2 of Appendix 3 illustrates the kinds of justificatory responses given. For each experiment, justifications associated with the critical pairs 12, 34 and 56 are listed in that order (along with the corresponding identical pairs 11, 22 ... 66); justifications for pairs differing by two attributes follow, and lastly any "unclassifiable" responses. For Experiment 4, <u>all</u> of the justifications that occurred are listed (p. xii), since there were so few in total. Only a sample is given for Experiments 5 and 6, representative of the range offered (pp. xiii - xiv and xv). Finally, samples of the descriptions of individual stimuli elicited during the preliminary familiarisation period of Experiment 6 are added (pp. xvi - xvii).

Response rates

Response rates were generally low since justifications were not specifically requested. Of the 40 subjects in Experiment 4, only 15 offered some justificatory response, on a total of 27 trials. The inverted stimulus presentation of Experiment 5 provoked more comment: 14 of the 27 subjects responded, on a total of 48 trials. The children given the "pretraining" on verbal description in Experiment 6 were the most inclined to respond during pair-comparison, producing a justification on a total of 40 trials, with only one of the 14 subjects making no response.

Order of reference

Responses containing reference to more than one attribute were scarce. In Experiment 4, there was only one multiple response, mentioning Mouth and Nose (see Appendix 3.2, Example 22). Experiment 5 produced eight multiple responses, six referring to Mouth first (Nose second). Experiment 6 included seven multiple responses, three with Mouth first and four with Eyes first. These are so few as to be generally uninformative, although in the 16 multiple responses generated in all, it is perhaps worth noting that Nose was first-mentioned only once (Experiment 5).

In the familiarisation period of Experiment 6, however, all children were required to comment on all three attributes for each stimulus. Overall, the distribution of these multiple descriptions among the six possible orders of reference was:

Mouth first,	38 times	(24 M>E>N, 14> M>N>E)
Eyes first,	33 times	(13 E>M>N, 20 E>N>M)
Nose first,	13 times	(5 N>M>E, 8 N>E>M).

The favoured order of reference tended to be related to the particular face being described. When Mouth was curved (faces 1 and 2), Mouth came first in 24 of the 28 descriptions. When Eyes were curved (faces 3, 4), Eyes tended to be first mentioned (18 times) -- otherwise Mouth was first. Faces 5 and 6, with Nose curved, produced more variation of order: Nose came first 12 times, Eyes first 11 times and Mouth first 5 times. Generally, then, Nose did not receive first mention except sometimes when Nose was the only curved feature present.

In the earlier discussion of the children's pair-comparison performance, the possibility was considered that visual scanning patterns might contribute to the distribution of correct matches among the attributes (see the discussion of those results in relation to Ghent Braine's work: pp. 133 - 135 and 145 -146). The question was whether the subjects were treating Mouth as a "focal feature" and were beginning inspection of the stimulus figures at Mouth, or whether they were following a top-down order of inspection.

When the stimuli were inverted (Experiment 5), both scanning strategies would have led the children to inspect the stimulus attributes in the order M>N>E (top-middle-bottom). The data on order of mention of the attributes in multiple justificatory responses, however, do not particularly correspond with such a pattern. Although in Experiment 5 only seven multiple responses were produced altogether, four of these placed the bottommost and presumedly non-focal attribute, Eyes, first.

With the stimuli in conventional orientation, attention to the focal feature first would have led to the scanning order M>N>E (bottom-up) or M>E>N (bottom focal feature first, then reverting to top-down); while a straightforward top-down pattern would have produced the order E>N>M. The descriptions given during the familiarisation period of Experiment 6 favoured the orders M>E>N and E>N>M about equally, with M>N>E and E>M>N next. The order in which the attributes were referred to thus provides no clear evidence in support of either visual scanning strategy. If order of reference can be taken as reflecting order of visual inspection, it seems rather as though with this stimulus set the children were attracted first to whichever of the attributes appeared in "unusual" -- i.e., curved -- form.

Frequency of reference to each attribute

Table 36 shows how the references were distributed among the three attributes. Eyes consistently received the fewest references. Mouth and Nose were referred to equally often in Experiments 5 and 6, but Mouth was the most commonly noted in Experiment 4. Interestingly, these frequency data do not exactly parallel the corresponding patterns of correct matches observed in the main tasks (see

TABLE 36

Distribution of references among the attributes (percentages of total references for each experiment)

	Mouth	Eyes	Nose	Total refs. *
Experiment 4	50	19	31	26
Experiment 5	38	20	42	50
Experiment 6	37	26	37	46

* Excludes "unclassifiable" responses

Table 27, p. 158), in that Nose featured considerably more prominently in the justifications than would have been anticipated on the basis of its rather poor pair-comparison results.

There is, however, a close parallel between mention of an attribute and its appearance in curved rather than neutral form: the child usually indicated any curved feature present, but seldom indicated the other, straight features. The strength of this pattern of noting just curved features is highlighted in Table 37. For clarity, only the figures for identical and single-attribute-different pairs are shown, but the same pattern was evident also for pairs differing by two attributes. Nor was it simply that the children were remarking on the one attribute whose variants marked the <u>difference</u> in a face pair, for this pattern of noting only the single curved attribute still obtained with <u>identical</u> pairs. In fact, occasions on which only a neutral feature was referenced were generally associated with wrong "same" pair-comparison judgments -- see, for example, Appendix 3.2, nos. 17, 31, 35.

<u>Types of reference</u>

The references to each attribute were classified by type following the system outlined on pp. 181 ff. For this second group of experiments, however, an extra category was felt to be necessary. A considerable number of responses now consisted of the child's remarking that a line (or an attribute, or a bit, etc.) was "going that

TABLE 37

Distribution of references among the attributes in relation to the structure of the stimuli (raw frequencies)

CURVED ATTRIBUTE	Number of references to attributes:								
	Mouth	Eyes	Nose	Mouth	Eyes	Nose	Mouth	Eyes	Nose
curved M (faces 1 & 2)	9	–	–	10	2	1	8	1	1
curved E (faces 3 & 4)	–	4	–	1	2	1	–	6	–
curved N (faces 5 & 6)	1	–	6	2	–	7	1	1	10
	EXPERIMENT 4			EXPERIMENT 5			EXPERIMENT 6		

Note:

Circled cells are where one would expect the references to occur if justifications have to do with remarking on curved rather than straight forms of attributes way" on one face and "going that way" (i.e., a <u>different</u> way) on the other face of a pair, as the child pointed to the specific aspect of difference to show exactly what he meant. While this did not capture the distinction between variants in language in the way that the usual Feature-level or Face-level responses did (i.e., by overtly saying "curved" vs "straight", "up" vs "down", "happy" vs "sad", etc.), the child was nonetheless clearly differentiating the levels. The undifferentiated Point/Name category seemed inadequate for such instances, and a new differentiated Point-level category was therefore added. In "hierarchical" order (see pp. 184 - 185), the categories were now:

<u>Face-level</u>: Child linguistically differentiates levels but goes beyond purely physical characteristics of the attribute to characterise the whole face by state or action. Examples: "happy" vs "sad", "sleeping" vs "awake", "looking up" vs "looking down".

<u>Feature-level</u>: Child linguistically differentiates levels in terms of objective physical structure of attribute. Examples: "curve", "straight", "going up", "curling down", "going along the way", "bent", etc. Also figurative descriptions such as "like a moon", "like a (ba)nana" (= crescent-shaped), "like a 'one'" (= straight line, as in a written number "1").

<u>Point-level</u>: Child differentiates levels but not fully in language. He points to the appropriate variants and says that one is "going one way" and the other is "going

the other way", or one is "like that" and the other is "not like that", etc., while tracing the respective outlines with his finger.

Point/Name: Child does not explicitly capture the distinction between levels in language. He simply points to or names an attribute to indicate that it is the same or not the same. Verbal examples: "Mouths the same"; "the nose goes a different way".

Other: Unclassifiable responses. Appendix 3.2 gives several examples.

In all other respects, the coding proceeded as described for the first two experiments. Notice, however, that now there are three levels of each attribute to be differentiated. It was relatively easy for the children to distinguish a curved from a neutral variant in words like "curly" vs "straight". But to distinguish the two curved variants or a single attribute from each other sometimes proved more difficult, for this involved both noting the curve and finding some way of specifying its direction. This is presumably why the new Point-level kind of reference came into play. When a child gave an ambiguous reference, such as that both Noses on pair 56 were "bent", the Experimenter pressed for clarification. If the child then seemed to be aware of the distinction and to be trying but failing to describe it verbally, the Experimenter offered terminology or suggested that he point his finger or face in the direction indicated by the curve. In cases where the distinction remained unclarified, the reference was classed as undifferentiated Point/Name despite the occurrence of a physical-descriptor term like "curve", since the critical difference of levels had not been specified.

Table 38 shows how the references were distributed by type for each attribute during the three pair-comparison tasks. Raw frequencies are presented, since numbers were so small. The corresponding figures for the descriptions given during the prior familiarisation period of Experiment 6 have been added.

Taking just the pair-comparison justifications, it is obvious that once again Mouth was the only attribute with which Face-level descriptions were associated; otherwise, Mouth was simply pointed to or named. Eyes references fell mostly in the Point/Name class with some Feature-level description, while Nose references were a mix of Point/Name, Point-level and Feature-level.

The pattern of descriptions from the familiarisation part of Experiment 6 is quite strikingly different. Point/Name references were more or less precluded here, since the Experimenter pressed for verbal description except in one or two odd cases where the child either seemed upset by the questioning or had already amply demonstrated his ability to describe the attribute's three variants distinctively. What is surprising is the large proportion of full Feature-level descriptions as opposed to the less explicit Point-level type, and the remarkably high proportion of Face-level references accruing to Eyes. As can be

TABLE 38

Frequencies of references of various types for each attribute

ATTRIBUTE REFERRED TO		TYPE OF REFERENCE			
		POINT/ NAME	POINT-LEVEL	FEATURE-LEVEL	FACE-LEVEL
Experiment 4 (conventional)	MOUTH	1	0	0	12
	EYES	3	1	1	0
	NOSE	4	4	0	0
Experiment 6 (familiar)	MOUTH	8	4	2	5
	EYES	7	0	3	0
	NOSE	8	6	7	0
Experiment 5 (inverted)	MOUTH	6	1	4	6
	EYES	6	2	4	0
	NOSE	11	4	2	0
Experiment 6 descriptions during familiarisation	MOUTH	3	2	47	32
	EYES	3	3	37	41
	NOSE	5	18	61	0

seen from the examples in Appendix 3.2, most of these were of the type "He's sleeping", "He's opened his eyes", "He's looking up". The "sleeping" or "eyes shut" kinds occurred especially with reference to the straight-line Eyes variant.

The child often ran into difficulty over Eyes descriptions here, however. What commonly happened was that he began by associating the neutral variant with *sleeping*. If on his first encounter with a curved variant he then captured the contrast with neutral by saying *wakened*, he had a problem when the second curved variant appeared and he had to find a third contrasting term! In such a situation, the child tended to "revise his strategy", using *sleeping* and *wakened* subsequently for the two curved variants and changing to *straight* for the neutral one, or switching to Feature-level description for all three variants. This could explain why Eyes did not later attract Face-level descriptions during pair-comparison proper -- the child had already found Face-level descriptions for Eyes led to confusion.

Once again, there was some relationship between the type of reference given and the physical structure of the stimulus it was connected with. For Mouth, most of the Face-level references were attached to faces 1 and 2 -- the critical Mouth pair, having curved variants. Similarly, almost all of the Point-level references to Nose occurred with faces 5 and 6. Nose on the other faces could be easily described as "straight" vs "curved" (etc.), whereas presumably the necessity of specifying direction of curve with 5 and 6 led the children to resort to the Point-level method. In keeping with what has already been said about the incidence of Face-level references to Eyes and the tendency here to switch later into Feature-level, there was no particular correspondence of type of reference to type of variant for Eyes. Where one might have expected faces 3 and 4 to account for most of the Face-level references, in practice no such relationship held.

Summary and discussion

In pair-comparison performance in this section of the study, Mouth was by far the most often correctly discriminated of the three attributes. Yet Mouth did not appear especially prominently in the children's justifications for their judgments. The frequency with which each attribute was referenced, and the order of reference among the attributes in multiple responses, seem instead to have been linked to the type of stimulus pair with which the reference was associated -- that is, to whichever of the attributes appeared in its curved form.

The descriptions elicited during the familiarisation period of Experiment 6 demonstrate that the children were well able to perceive and discriminate the variants of all three attributes when the stimuli were viewed individually, and stand in marked contrast with their subsequent pair-comparison performance and justifications.

There were no outright erroneous descriptions of the features of the second set of stimuli. Where there were failures here, they were either of ambiguity or of omission -- e.g., saying "the nose is curly" for both faces 5 and 6 without specifying the opposite directions of curve, or justifying a wrong "same" decision by indicating an identical attribute but not mentioning the attribute(s) that differed. But such "inadequate" descriptions of the stimuli were comparatively infrequent, and there is little evidence from the justificatory responses of a failure to detect Nose or Eyes differences that would correspond with the proportion of confusion errors made on these two attributes relative to Mouth in the pair-comparison judgments.

The finding that in Experiments 5 and 6 references to Nose were at least as frequent as references to the more dominant attribute Mouth is striking in this regard, and suggests that the children were aware of the Nose although they often called faces actually differing in Nose the same. Moreover, references to Nose occurred despite the fact that this attribute posed a lexical problem paralleling that of Shape in the previous experiments, in that it was difficult for young subjects not yet able to apply the terms "right" and "left" to capture verbally the distinction between its two curved variants. In the case of Nose, however, the children were able to overcome the problem by resorting to the Point-level kind of reference.

<u>Understanding of the word "same" in the context of the test-question</u>

There is some suggestion in the justificatory responses that the curved Nose forms appeared "odd" to several of the subjects. The large majority of the unclassifiable responses that occurred were associated specifically with Nose -- see Appendix 3.2, Examples 26 and 49-55. (This was also remarked on earlier: p.125 footnote, and pp. 138 - 139.) It is clear from the appended examples that a straight Nose was judged to be the same and a curved Nose to be not the same -- but (not) the same as <u>what</u>? The child seemed to be comparing each test face separately against the stimulus set as a whole, i.e., to be answering the question "Are they the same <u>as the others</u>?" instead of "Are they the same <u>as each other</u>?". And in that comparison, evidently, faces containing a curved Nose were seen as "out of place". Similar responses emerged in the earlier pair-comparison task of Experiment 3 with the first stimulus set (see Appendix 3.1, Examples 42, 44, 46), although there they were not tied to any one of the attributes in particular.

Children who made these kinds of responses appeared to be assessing class membership and hence taking the test-question to refer to "same kind of entity" (as the others) rather than to "perfect identity of features" (between a pair). This reflects a syntactic/pragmatic misunderstanding of the question posed to them, i.e., of the task requirements. It is, however, quite distinct from the kind of immature appreciation of the semantic components of the word "same" itself that is suggested by the semantic-feature theory type of explanation for children's confusion errors in sameness judgments(cf. pp. 5 - 7).

In fact, a fair proportion of the justificatory responses displayed a remarkably well-developed sense of the meaning of "same", incorporating such subtleties as the notions that entities can be more or less alike and that similarity without perfect identity entails difference too. Expressions such as "they're <u>nearly</u> the same (but) ... ", "<u>not quite</u> the same", "<u>not at all</u> the same", "the <u>very</u> same", and "they're the same but they're different" were by no means uncommon. (For illustration, see Appendix 3.1, p. v, Mouth-matcher on pairs 14 and 15; 3.1, Examples 2, 3, 5, 11-13, 15, 22, 23, 28, 32, 36, 39, 41; 3.2, Examples 29, 37.) These point to a degree of linguistic and perceptual competence beyond that which one would have supposed from main-task performance alone, if one considered simply the relatively large proportions of confusion errors without the justificatory responses.

The children's use of the word "but" in justifications on trials involving non-identical face pairs is interesting in this regard. Clauses beginning with "but" occurred in two main types of context. In one, the child made a <u>successful</u> discrimination, saying something like "They're not the same but (a particular indicated feature) is the same". This reflects competent, adultlike performance.

In the other type, the child was scored as making a confusion error following his decision "They're the same ... "; however, he then added something like " ... but (a particular feature) is not the same" -- or, equally, "(A particular feature) is not the same but they're the same faces". Adult subjects never gave this sort of response; if the stimuli differed in any respect, the adult judgment was "They're not the same because (a particular feature) is not the same". In such instances, the Experimenter often asked the child additional questions to the effect: given the presence of some difference, could the pictures still be exactly the same? Almost invariably, the child remained content that they were the same while openly acknowledging that a specified part differed -- and despite the fact that one might have expected him to reverse his judgment on the grounds that the extra questioning implied the Experimenter thought his first answer was wrong. (Appendix 3 contains many examples of such responses. See especially 3.1, Example 29; also 3.1, p. v, Mouth-matcher; 3.2, Examples 16, 18, 36, etc.)

In the pair-comparison task of Experiment 3, with the first stimulus set, a total of 78 of the confused trials received justificatory responses. In more than 60% of these, although judging the pictures to be the same, the child correctly noted that a particular, specified attribute differed.* This is a substantial proportion, and is persuasive evidence against the contention that confusion errors were due to failure to perceive an attribute or failure to detect the difference between its levels. Furthermore, the children's generally sophisticated use of the word "same" exhibited elsewhere in the justifications, together with the non-random nature of the distribution of confusion errors among the attributes, help to eliminate immature understanding of "same" (à la semantic-feature theory) as a major contributor to the confusion errors. The argument now seems more convincing that, at least with the first stimulus set of schematic faces, confusion errors arose because the children chose deliberately to ignore certain differences.

The evidence is less weighty with respect to the second set of stimuli, used in Experiments 4 - 6, but mainly because of the smaller numbers of justificatory responses produced altogether, not many of which fell on trials receiving a wrong "same" judgment. Nevertheless, in more than half of these few instances, the child again explicitly noted some aspect of difference along with his assertion that the pictures were the same.

In Experiment 3, about three-quarters of these justifications were to the effect that Mouth was the same but Shape or Eyes was not. In Experiments 4 - 6, this kind of justification fell almost exclusively on pair 56, and was to the effect "They're

---

* This phenomenon has also been recorded by other investigators, e.g., Oléron (1962), Ricciuti (1963, cited in Gibson, 1969), Taylor & Wales, (1970), Vurpillot & Moal (1970), Kemler & Smith (1979).

the same but the noses are going different ways" (without explicitly mentioning the identical attributes). Thus it appears that Shape/Eyes in the first stimulus set, and Nose in the second, were deliberately being treated as irrelevant for the task in hand, and that Mouth was actively selected as a major criterion of sameness/difference, on at least a considerable proportion of the trials.

Perhaps there is some connection here with the incidence of Face-level references in the justifications? For, recall, Face-level references were associated almost exclusively with Mouth throughout all of the main tasks. These sorts of descriptions involved characterisation of a stimulus figure by reference to just the one of its attributes; that is, the entire figure was globally identified (as, say, a "happy face" or a "sad face") in a manner that depended solely on the form of the Mouth. Assuming the justificatory responses reflected something of the way the subjects perceived the stimuli and of their underlying decision processes, then what the children were doing here could be taken as amounting to classification of the entire stimulus set on the basis of a single attribute. Thus the first set could be partitioned in two, according to whether the Mouth was curved or straight, while the second set divided into three, faces with Mouth curved up (happy), curved down (sad) or neutral (straight). Such a notion is appealing because it provides a convenient, economical way of labelling stimuli -- of both linguistic and perceptual-cognitive encoding -- that would aid the child's decision-making by reducing information load.

If the children were indeed engaging in such stimulus grouping, the implication is that -- like the subjects noted earlier (pp. 211 -212) who seemed to treat the curved Nose as not the same in relation to the rest of the stimulus set -- they understood the term "same" in the sense not of "identical in all respects" but of "belonging to the same class of entity". While such an interpretation would be quite legitimate in many circumstances and is common enough adult usage, in the present stimulus and task context it is considered inappropriate, since adult subjects always operated on the basis of number of identical features, not on a classification principle.

Classification, discrimination and sameness judgment are, of course, closely-related activities. Obviously, classification behaviour is at least partly dependent on perceived similarities and dissimilarities among the items being categorised. But the reverse can also be true: a subject's sameness/difference judgments may be influenced by the way he sorts the stimuli into categories (see, e.g., Garner, 1966; Gregson, 1975; pp. 221-223).

This latter effect has been neatly demonstrated by Tversky (1977). He notes that, in order to optimise information capacity and processing, stimulus classification must involve maximising both the similarity of members within groups and the dissimilarity between groups. The posited relation between similarity and grouping he calls the "diagnosticity hypothesis". "Diagnosticity" refers to the importance that a given stimulus feature has for classification purposes, and it varies according to the structure of the stimulus set within which the feature appears. Tversky obtained support for his diagnosticity hypothesis through experiments using stimulus sets composed either of so-called "visual features" (schematic faces) or of "semantic" material (names of countries). Adult subjects were given either a matching-from-sample type task or were instructed to sort the stimuli. The resulting match-choices and groupings, although independently obtained, both followed the same patterns and changed in tandem with minor alterations to the stimulus context.

Tversky's stimulus sets contained no totally identical pairs, so that subjects given the matching task were obliged to respond in the "similar in some respect(s)" manner. What is important is his finding that, under those circumstances, the subjects' choice in the similarity judgment was not founded on sheer physical identity. Their criterion of similarity was at least partly subjectively determined, and it paralleled the criterion selected (by other subjects) as the basis of classification -- even when the subjects were adults. This lends weight to the argument that the children in the present study who produced a dominant-attribute matching pattern with confusion errors on non-dominant attributes may have been partly operating by what amounted to a classification strategy.

The final part of this study moves on to test an hypothesis about the basis on which the children may have been classifying/judging sameness among the schematic faces.

First, however, a supplementary experiment was conducted on the detection-versus-selection of attributes issue.

EXPERIMENT 7

Matching Game

In none of the pair-comparison tasks so far presented was the precise sense in which the subjects were expected to take the word <u>same</u> made overtly clear. That is, there was nothing explicit in the task format or the Experimenter's questions to tell the child that he should judge in terms of sameness of <u>all</u> features rather than of similarity in <u>some</u> way or same <u>kind</u> of thing. (The fact that adults always took the first interpretation is irrelevant to this point.) The child had no way of knowing when his answer was wrong, by the adult's expectations, since no specific feedback was provided about the (in)correctness of responses. Instead, he received general verbal encouragment at random intervals -- which, in the case of an incorrect matcher, may have had the effect of encouraging him to keep on making the same kind of confusion errors!

The familiarisation period of Experiment 6 was designed to induce subjects to attend to all of the stimulus attributes and to ensure discriminability of variants. The attempt was successful to the extents that the children's preliminary descriptions of individual faces did distinguish all stimulus features with fair precision and subsequent pair-comparison performance was generally superior to that observed in

Experiment 4, when the faces had been unfamiliar to the subjects. It failed in that even after practice the children typically had to be prodded into giving complete, unambiguous stimulus descriptions (clearly shown in the samples in Appendix 3.2, pp. xvi - xvii, by the number of appearances of "(Q)", standing for "further questioning by the Experimenter"), and in that Eyes and Nose continued to be confused in pair-comparison significantly more often than Mouth.

Following that "failure", Experiment 7 was added as a supplement, out of curiosity as to whether young children <u>could ever</u> operate on the basis of complete identity of all the stimulus features, and could ever volunteer complete stimulus descriptions. The aim was to promote perfect matches and full descriptions, if they could indeed be achieved, by establishing for the child that partial description and partial similarity would not suffice -- by providing him with explicit feedback that each response was correct, wrong, or inadequate according to the adult's expectations.

The idea for the experimental task stemmed from the research of Glucksberg, Krauss and Weisberg (1966) and Glucksberg and Krauss (1967) on referential communication between children, and roughly followed their model. A game-like situation was set up in which one child had to describe a stimulus figure to another child in such a way that the second child could select that same figure from an array with only his partner's description to go on.

Materials

The first two sets of schematic faces (Appendix 1, Figures 1 and 3) were each employed, with separate groups of subjects. The faces of the first set were mounted on square white cards as before, but with the surrounding square frame omitted. Preliminary training was given with an assortment from the stick-figures and geometric forms already used for practice trials in the previous experiments. Three copies of each stimulus set were required: the two subjects per game each had on view a complete array of the appropriate set, and the first child had an additional copy of the set stacked in a deck.

Subjects

The first stimulus set was presented to a final sample (after drop-outs) of 29 pairs of children. Fourteen pairs were in a nursery unit in St. Andrews and came from a variety of home backgrounds, while the remaining 15 were in a community-centre playgroup in Oxford that pooled mainly from working-class families. Their ages ranged from 3;10 to 5;2, with a mean of 4;7 years. Ten more pairs, all from the Oxford community-playgroup, were given the second stimulus set (after which the attempt with the second set was abandoned). Their ages were 4;2 to 5;3, mean 4;8 years.

Procedure

A table was set up in a corner of the nursery or playgroup room, with a chair at each of three sides. The children were invited in pairs to come and play a game with pictures. The two children sat facing each other on opposite sides of the table, with the Experimenter occupying the middle chair. A cardboard screen was erected across the middle of the table so that neither child could see the other but the Experimenter could see both.

The Experimenter laid on the table before each child an array of all the stimuli of the appropriate set. The eight faces of the first set were arranged in a 3 x 3 square with a space in the centre bottom row. The six faces of the second set were displayed in two rows of three. Within a subject-pair, the locations of stimuli in the array were the same for both children and remained fixed throughout their game. The cards were re-ordered for the next subject-pair by shuffling. The two subjects were told that they each had the same group of pictures in front of them but that they would not be able to see each other's.

On each trial, the Experimenter shuffled the third deck of copies of the appropriate stimulus set and handed the deck face-down to the first child, having surreptitiously looked to see which card was on top of the stack since this would be the test stimulus on that trial. The instructions were as follows, with "C1" and "C2" substituted for the respective names of the first and second child. (Each child heard his partner's instructions as well as his own.)

<u>To C1</u>: Turn over the top card and have a good look at it. Don't show it to C2 -- you must keep it hidden. C2 is going to try to find which one of the pictures you are looking at. He's got to find the <u>only</u> one that is <u>just the same</u> as your picture. You have to help C2 find the right card. Tell him as much as you can about your picture. Remember C2 can't see your picture, so you have to tell him <u>everything</u> about it so he can find the one that's <u>just the same</u>. (Experimenter pointed to the card in C1's array that matched the test card he was holding.)

<u>To C2</u>: C1 is looking at a picture that's <u>just the same</u> as one of yours. He'll tell you about his picture, and you have to find which one of all of these (Experimenter indicated C2's array generally) is <u>just the same</u> as C1's. When you think you've found it, hold it up like this (E. demonstrated holding up card above the dividing screen) so C1 can see if it really is <u>just the same</u> as his. If you can't find the right one on your first go, then you can ask C1 some more about what his picture looks like.

<u>To C1</u>: When C2 holds up a picture, you have to tell him if it's the right one. Look carefully at his picture, and look carefully at yours, to see if they <u>really are just the same as each other</u>. Then tell him right or wrong.

<u>To both</u>: Let's see if you can get them <u>just the same as each other</u>.

The Experimenter demonstrated the entire sequence, then gave the children several practice trials with the training figures until they seemed comfortable with the procedure. At least six practice trials were needed, largely because C1-subjects persisted in trying to show C2 the card instead of describing it. C1's earliest attempts at description were generally inadequate through omission of stimulus features, but improved as C1 learned that C2 could not find the correct match. C2's choices as match to the description, and C1's decision as to whether or not the choice was indeed a correct match, were satisfactory from the start of the practice. Following practice, the children received three or four trials with the test stimulus faces. Although the children were enthusiastic about the game, it was decided to limit the number of test faces to only a portion of the full set because of the length of time taken for each trial and the large amount of practice required beforehand.

The matches that C2 chose to C1's descriptions could be included in the data only when C1 had on that trial given a satisfactory description that uniquely distinguished the test stimulus from all the others in the set. The rest of the time, C2 served as unwitting "Experimenter's accomplice" to induce C1 to amend his so far problematic description. That is, when C1 gave a description that was false, ambiguous or omitted a feature, the Experimenter "helped" C2 to find a face that matched the description as far as it went but differed from the test stimulus in the unsatisfactorily-described attribute(s). This mismatch was then held up for C1 to judge whether or not it was the same as his test card. If at this point C1 gave a wrong "same" judgment, the Experimenter indicated that and where an error had occurred, and asked C1 to try again to describe the test face. A total of four attempts was permitted before the trial was abandoned, but only the first attempt was scored in the data. No C1 ever gave a wrong "not the same" judgment when an identical match was offered. If C2 chose a mismatch to a satisfactory description, C1 showed him the test card, and C2 then tried to find a match to the visual stimulus.

No hypotheses were tested, no predictions made; the experiment sought simply to "see what happens".

Results

This game yielded data concerning three types of performance:

(1) C2's choices as matches to C1's descriptions (where the latter were satisfactory); these were equivalent to matching-from-sample choices (cf. Experiment 1), except that the "standards" were described instead of presented for visual inspection.

(2) C1's decisions as to whether or not C2's choices were correct matches to the test stimuli; these amounted to (visual) pair-comparison judgments.

(3) C1's verbal descriptions of the stimuli; these corresponded to the justificatory responses and familiarisation-period descriptions from the previous experiments.

The first of these can be quickly dealt with. A total of 109 trials was presented to the subject-pairs who received the first stimulus set. On 74 of these, the C1-subject produced a satisfactory stimulus description. And in only three of those 74 instances did the C2-subject select a mismatch. This was an astounding achievement -- compare with the matching-from-sample performance of Experiment 1 (Tables 7 and 9, pp. 54 & 56), where the highest proportion of perfect matches attained by any subgroup of subjects was 56% (scored by Group I with Framed stimuli) and the overall success rate was a mere 39%!

With regard to the C1-subjects' "pair-comparison" performance, it has already been noted that no wrong "not the same" judgments were given to actually identical pairs. With the first stimulus set, C2-subjects presented a total of 38 mismatches for C1s to compare with the test-stimulus (109 trials in all, minus the 71 correct matches-to-satisfactory-description, considering first attempts only). Eight of these were wrongly judged to be the same as test figures. This success rate of 79% is marginally better than the corresponding pair-comparison results of Experiment 3, which yielded 72% correct overall (Table 16, p. 85).

All three of the C2s' mismatches, and all eight of the C1s' wrong "same" judgments, involved differences of just one attribute, and never of Mouth. That is, the only confusions that occurred were of Shape or Eyes, and Mouth was universally correct.

The case was rather different for the second stimulus set. Here, a total of 37 test trials was presented, 27 of them producing satisfactory descriptions. Only two-thirds of these 27 "standards" were correctly matched by the C2-subjects. Nineteen mismatches were thus returned to the C1s for pair-comparison. This time, some of the mismatches differed from the test figure by <u>two</u> attributes, and while Nose was almost always among the differing attributes, Eyes and Mouth were each confused on occasion. The C1s made correct "not the same" judgments on only 11 of these (58%). In this phase, Mouth was again universally correctly discriminated, and Nose most often confused.

On the surface, these results with the second stimulus set appear rather poor, but there is no way of directly comparing them with earlier performance, e.g., in Experiment 4. For one thing, matching-from-sample was never previously attempted with this set. For another, the earlier pair-comparison results were not examined in terms of percentage of correctly-judged pairs overall; the focus had been on successful discriminations of the three critical pairs separately, differing by Mouth, or Eyes, or Nose. In Experiment 4, Mouth was discerned on three-quarters of its critical trials, Eyes on half, and Nose on only a quarter. In this light, performance in Experiment 7 is seen to be superior, since most of the problems here occurred with Eyes and Nose again, yet with better than 50% success even on Nose.

The C1-subjects' descriptions of the test stimuli were examined in the same way as were the justificatory responses of the previous experiments. Samples are not necessary since they followed the pattern of the familiarisation-period descriptions in Experiment 6 (see Appendix 3.2, pp. xvi - xvii), except in that the Experimenter did not intervene with questions and prompting. Although the C1s received feedback in the form of the C2's choices of match or mismatch and could then amend their descriptions, only the first attempt at each description counted here.

<u>Order of reference to attributes</u>. With the first stimulus set, the orders of mention E>M>S and S>E>M were equally popular, with M>E>S next. Thus there was no preference for any one attribute over another in first place. The second stimulus set elicited references mostly in the

single order E>N>M (corresponding to top-down visual sequence). These patterns disagree with those found earlier in Experiments 2 - 6, where either Mouth tended to come first in the sequence, or else the order depended on the particular stimulus being described.

Type and adequacy of reference to attributes. References to each attribute were classified as Face-level, Feature-level, Unclear, or Wrong. To be classed as Face- or Feature-level, the reference had to be accurate. Wrong references were false and misleading ones, e.g., calling black Eyes "white". Unclear references were not erroneous descriptions but failed to differentiate an attribute's levels; they were usually ambiguous (e.g., calling a curved Nose "crooked" without specifying direction of curve), or so idiosyncratic as to be largely meaningless, even though the child himself might have been successfully differentiating attribute-levels according to his own private system (e.g., calling Eyes "red" when there was no way the listener could tell whether that meant black or white). Unclear references also covered total omission of an attribute, though this happened very rarely. Point/Name and Point-level references were of course precluded here, when the C2 could not see the C1's test stimulus.

Table 39 shows the proportions of the various types of reference for each attribute. Although percentages have been used for ease of comparison, it should be noted that with the second stimulus set these were based on a total of only 37

TABLE 39 (Experiment 7)

Percentages of references of each type for each attribute

ATTRIBUTE REFERRED TO		Face-level	Feature-level	Unclear	Wrong	Total no. references
		TYPE OF REFERENCE				
First stimulus set	MOUTH	61	27	7	6	109
	EYES	0	88	7	5	109
	SHAPE	1	80	15	5	109
Second stimulus set	MOUTH	41	49	11	0	37
	EYES	14	73	11	3	37
	NOSE	0	73	28	0	37

references, so that the "3%" entered in the Wrong column actually represents only one Wrong description in all.

Predictably, Mouth was as usual associated with a high proportion of Face-level references, especially with the first stimulus set. The percentage is lower for the second set because here the happy/sad dichotomy no longer sufficed, and the neutral variant now tended to attract Feature-level descriptions ("straight") while "happy" and "sad" were applied to the two curved variants. False descriptions were equally distributed among all the attributes in the first set and were uniformly rare; in the second set, only the one false reference occurred altogether. The two attributes that have been shown before to be the most difficult for the children to describe verbally, Shape and Nose, were those that caught the largest share of Unclear responses, that is, ambiguous or idiosyncratic descriptions.

It was problems over Nose and Eyes that led to the abandonment of the game with the second stimulus set after only 10 subject-pairs had been tested. The C1s were struggling over their descriptions and became frustrated when they could not find the right words. Even when they did manage what seemed to the Experimenter to be a satisfactory description, it might turn out in practice to have been ambiguous, eliciting a mismatch from the C2. Thus it was one matter for the C1 to call both forms of curved Eyes "round" -- this simply provided insufficient information and was a fault on the C1's part. It was another matter when C1 called the curved-up Eyes "rounded up the way", a seemingly fair description but one to which C2 responded with the curved-down variant, showing that for him at least the description was unclear. Similarly, "sleeping" might refer to the curved-up Eyes variant for one child but to the neutral form in his partner's system (recall the parallel difficulties the children in Experiment 6 experienced within their own systems when they began with Face-level descriptions of Eyes and subsequently had to adjust their terminology or switch to Feature-level reference). As each child persistently failed his partner, the trials took so long and the subjects became so exasperated that it was decided to discontinue this part of the experiment.

Adequacy of whole stimulus descriptions. The children's first attempts to describe each stimulus were examined in terms of whether they uniquely distinguished the test face from the rest of the set, or led to choice of a mismatch because false information was conveyed, or left the choice open because precise information was lacking. With the first stimulus set, 68% of the test faces were described completely and accurately, 20% imprecisely in some way, and 12% falsely in some way. The figures for the second set were 73% complete and accurate, 24% imprecise, 3% (one test face) false. (These proportions were based on the Experimenter's assessment of the C1s' descriptions, not the C2s' reactions.) These percentages compare favourably with the success rates achieved in the matching-from-sample and pair-comparison decisions from the earlier experiments. Notice that failure was associated considerably more often with inadequacy/ambiguity of description than with outright false description. In a number of the ambiguous cases, the C1 was actually clearly able himself to distinguish all three stimulus attributes but could not convey the distinctions in words. C1s frequently tried to use a (correct) Point-level reference, showing the stimulus to the Experimenter, but of course this did not help the C2.

Summary and conclusions

Experiment 7 added little new to the findings from the previous experiments in terms of the kinds of responses given. As before, Mouth proved the easiest attribute to describe clearly and accurately, and the form of Mouth was taken to characterise a state or action of the whole face. Shape and Nose were the most difficult for the children to describe verbally, and their precise form was sometimes left ambiguous. Recall, however, that in the familiarisation period of Experiment 6, the children rarely spontaneously produced a full stimulus description; one or more features were typically omitted and were mentioned only after considerable prompting. In contrast, the present game situation generated complete, accurate descriptions on the first attempt (i.e., before the child received feedback in the form of his partner's match-choice) in about 70% of cases. Thus while manner of referring to stimulus features was much as before, there was a general improvement in adequacy of description in terms of number of attributes mentioned and accurate differentiation of variants.

The children's choices of match-to-described-standard and the pair-comparison judgments as to whether the choice was really the same as the "standard" likewise showed improvement over past performance. This was especially striking in the matches chosen to the test stimuli of the first set, which were perfectly correct in 96% of cases, as compared with the general success rate of less than 50% in Experiment 1.

In answer to the questions that Experiment 7 set out to investigate, then, the reply is that this task design, with its provision of specific feedback on (in)correctness of responses, did yield higher levels of performance than did the previous experiments where the children had not been penalised for responses based on partial identification of stimulus features. Yet performance still fell short of adult standards: unsatisfactory descriptions and incorrect matches and sameness judgments still occurred, even if less frequently than before.

It must be added, however, that the task and stimuli employed here were perhaps not optimal for assessing whether or not young children can ever attain perfect sameness performance.* Glucksberg and his colleagues, whose

---

* On an informal note, the "picture-lotto" games (resembling "bingo" but with pictures instead of numbers) commonly found in preschool centres provide a similar game, but with all stimuli visually presented and no verbal description needed. The task involves both matching-from-sample and pair-comparison type responses. Preschool children are quite capable of perfect performance in these games, successfully selecting matches and rejecting mismatches even when the differences are rather subtle. Many other "educational" toys and games that preschool children can handle competently also involve, in one way or another, the concept of sameness = perfect identity.

experimental paradigm the present game followed, found that their youngest subjects faired rather badly when performing such a task with novel stimuli (outlines of irregular shapes) for which they had no existing verbal labels. Like the present subjects, their "kindergarteners and first-graders 'point' -- that is, said things like 'it goes like this', while tracing the design with a finger" (Glucksberg & Krauss, 1967, p. 314), a means of communication unhelpful to a hidden partner. Their young subjects' stimulus descriptions tended to consist of a brief and idiosyncratic name that could not properly be called "descriptive" at all. Whereas adult subjects could find terms that both parties shared, the children adhered to their private, figurative references, unshared by the partner, and did not alter their idiosyncratic terminology even in the face of feedback that the partner could not find a match on the basis of such a "description" (Glucksberg et al., 1966; pp. 338, 341; Glucksberg & Krauss, 1967, p. 312; see also Flavell, 1975).

The present results agree on the incidence of Point-level and idiosyncratic references. In the present case, however, Unclear references that the C2-partner was not privy to were associated not with an entire stimulus figure but with a single attribute, Shape or Nose -- the other attributes were satisfactorily described almost 90% of the time. Since Shape and Nose were obviously genuinely difficult for the young subjects to describe verbally, less-than-perfect performance here probably had little to do with inability to discriminate variants perceptually and to take all features into account in sameness decisions. Thus face 5 of the second set (straight Eyes and Mouth, curve-right Nose) was satisfactorily described only half the time; but perfect performance, 100% satisfactory in descriptions, match-choices and pair-comparison judgments, was attained with face 1 -- which was the easiest of the set to distinguish, in terms like "happy with straight eyes and a straight nose".

It is possible, then, that under suitable task and stimulus circumstances, young children could attend to all features in their stimulus descriptions and decisions about sameness.

Finally, it should be noted that the kind of stimulus description given by the C1-subject in this Glucksberg et al. type of task can vary depending on the child's socio-economic background. Part-Descriptive encodings -- those describing physical properties of just part of the stimulus -- have been found to be used more by middle than by lower class children, while Whole-Inferential encodings -- those describing the whole stimulus figuratively -- were more common among lower than among middle class children (e.g., Heider et al., 1968; Johnston & Singleton, 1977; see Heider, Ref. Note 4 for the classification system). These findings are relevant to the present Experiment 7: half of the subjects given the first stimulus set here came from a largely working class population while the rest of the sample tended towards middle class; all subjects tested with the second stimulus set here were from the working class source. Since Feature-level references to attributes correspond to Part-Descriptive encodings and Face-level references follow the Whole-Inferential style, one might expect differences in type of reference according to the source of the subject. Although this was not examined here, the possibility should be borne in mind. (However, inspection of the results for subjects given the second stimulus set [working class] compared with those given the first set [mixed group] suggests little difference between the two -- see pp. 228-9.)

Part 5 returns the emphasis to the patterns of confusion errors among the attributes observed in the previous experiments, where the subjects appeared to be operating on the basis of partial similarity whether or not they were actually capable of discriminating confused features.

PART 5

CRITERIA FOR SAMENESS AMONG SCHEMATIC FACES

In all of the pair-comparison tasks conducted in this study, the majority of subjects were Mouth-matchers (or matched Mouth along with another attribute). Mouth also appeared prominently in the children's justifications for their sameness decisions -- even for subjects, like the Shape-matchers in Experiments 2 and 3, who did not show a Mouth-dominant matching pattern -- and references to Mouth were of a different type (Face-level) from references to other features. Throughout all of the preceding experiments, Eyes, in contrast, have consistently attracted a high proportion of confusion errors, whether the levels to be discriminated involved variation of colour within a fixed outline (black versus blank-white circles) or of orientation of the outline itself (mirror-image curves, up versus down).

The present section examines one possible reason for the occurrence of these consistent patterns of Mouth dominance and Eyes confusions. Two of the main contenders as explanations have been eliminated already: the "objective saliency hypothesis" -- that the children failed to detect certain variations because they were not visually salient by virtue of their objective, physical structure -- was ruled out in Experiments 4 to 6 through the device of equating the structural properties of all attributes and their variants; and the "location-of-gaze hypothesis" -- that the children were directing visual attention primarily to the bottom of the stimulus (Mouth) -- was disproved in Experiment 5 when the stimuli were presented upside-down so that Mouth appeared at the top. A third possibility, the "subjective weighting hypothesis", however, not only remained uncontested on the experimental evidence but actively received support from the children's justificatory responses, which demonstrated that subjects were frequently aware of the differences in non-dominant attributes' variants when they called unidentical stimuli the same.

This subjective weighting account posits that Mouth was deliberately selected as more important than the other attributes and that Eyes variations tended to be deliberately treated as irrelevant to the question of the sameness of these schematic faces. Then what was it about the stimuli or the attributes that led to the assignment of more weight to Mouth and less to Eyes? Is the mouth typically a criterial attribute, in the Gibsonian sense, for young children dealing with faces in general in such situations? Or was there rather something peculiar to the structure of the particular stimulus sets so far employed here, such that the patterns of matches and confusions observed up till now might not generalise to other designs of schematic or realistic face stimuli?

The clue comes from the justificatory responses. We have already noted that Face-level encoding could provide a shorthand way of stimulus grouping to which the children might resort in order to facilitate information-processing and decision-making. In the previous two sets of schematic faces, the form of the Mouth seemed especially to lend itself to Face-level description, much more so than the other attributes. It might therefore be that the pair-comparison judgments tended to be based primarily on Mouth simply because its particular structure here provided a more ready means of labelling and classifying the stimuli than did the other attributes (some of whose forms were extremely difficult for the children to label verbally).

But there is more to the story than that. References of the Face-level type were defined as those which went beyond the physical form of the feature itself, taking a single feature to characterise the whole face as being in some state or action. From the many examples that have been supplied, it will by now have become obvious that Mouth was not simply associated with a state or action, but with a particular kind. Almost universally in these sorts of responses, Mouth was taken to represent an affective facial gesture or mood: happy, sad, smiling, grumpy, etc. This immediately suggests the possibility that it might not have been Mouth per se that the children were attending to, but the affective significance of its form. In that case, Mouth could be regarded as a criterial attribute only in an incidental sense -- because in the particular stimulus sets used so far, Mouth chanced to be the only one of the attributes to be constructed in such a way as to encourage the ascription of affect.

EXPERIMENT 8

Pair-comparison with affectively-significant stimuli

It was now taken for granted that the children who made systematic confusion errors in the present study were deliberately focussing on only certain aspects of the schematic faces. Experiment 8 was designed to test the possibility that what they were attending to was facial expression. The hypothesis was that faces were being called the same if they could be taken as alike in affective expression, and were being called not the same if they expressed different affective messages. By this account, any affectively-significant feature(s) would appear to be treated as "criterial" while those features not interpretable in affective terms would tend to be ignored. This would explain the patterns of correct Mouth judgments and Eyes confusions consistently observed throughout the preceding pair-comparison tasks, for any affect expressed in the faces of the first and second stimulus sets was conveyed solely by Mouth. Alternatively, those results could be interpreted from the more traditional criterial-attribute standpoint: that Mouth has special significance in its own right and that, given roughly equal physical saliency and discriminability across attributes, Eyes would always incur relatively more confusion errors, regardless of any affective connotations. The affect hypothesis, however, received additional support from the children's verbal descriptions of the stimuli and justifications for their decisions (examined in Part 4).

To test the proposal that, when possible, the children would take (dis)similarity of affect as their criterion rather than automatically giving Mouth the most weight for its own sake, a new set of stimulus faces was required in which Mouth was not the only attribute whose distinctive levels were encodable in affective terms, but in which form and type of transformation were not so disparate <u>across</u> attributes that the "unequal physical saliency" objection could be raised again. The first condition was easily met: a person's face-shape or nose do not normally convey information about affect, but his eyes do. So the two critical attributes would be Mouth and Eyes. But the second stricture was harder to comply with. While the mouth expresses different emotions by wide muscle movements that alter its whole outline, changes in the eyes are more subtle. For continuity with the previous experiments, it was desirable to retain the curve-up/curve-down Mouth forms of the second stimulus set. The Eyes variants then also had to consist of single lines. The forms ultimately chosen played on the mobility of the forehead muscles, caricaturing lowered and furrowed eyebrows.

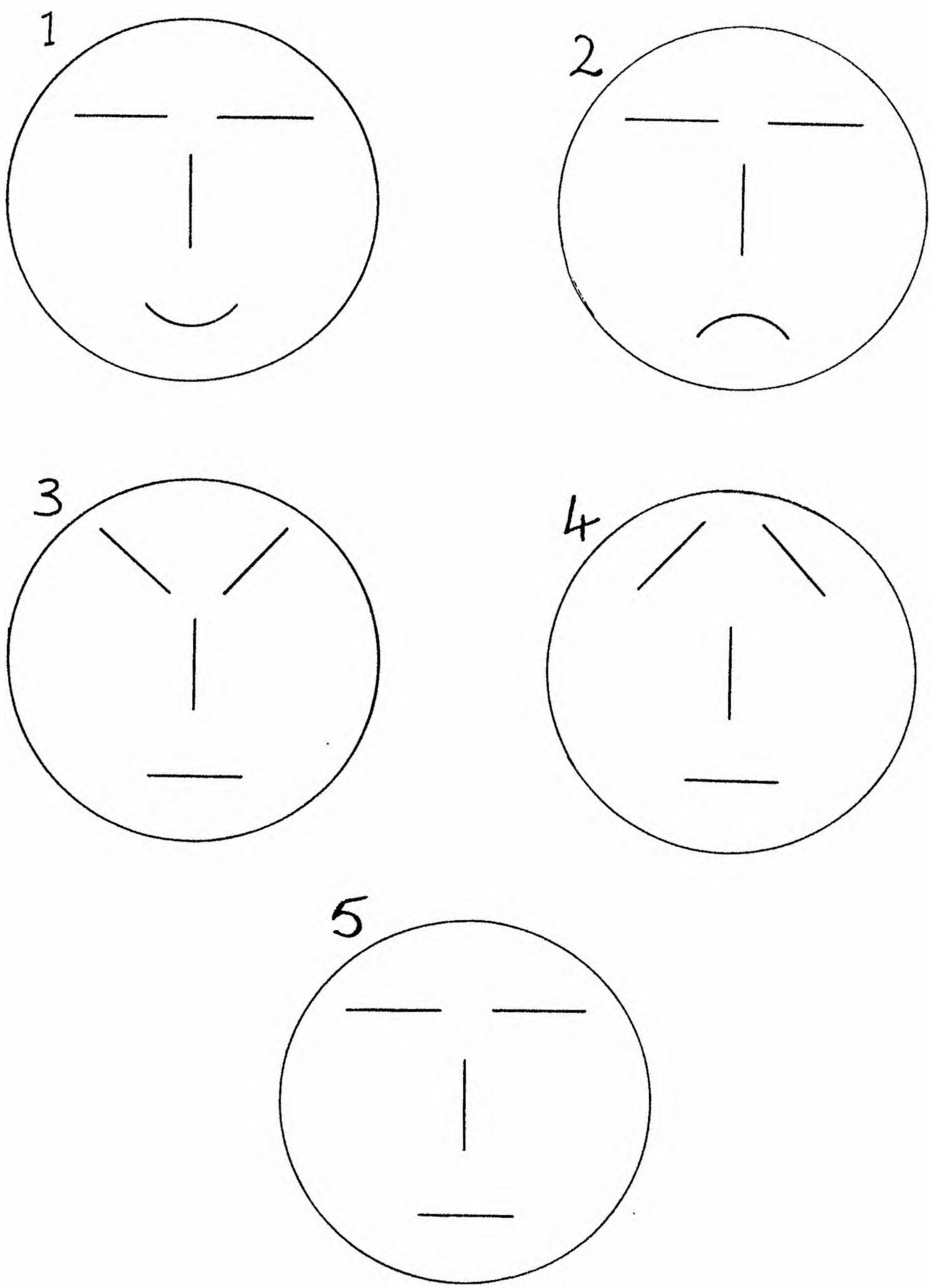
Materials

As much as possible was carried over from the second stimulus set (see pp. 115ff for its design). The original circle-template (Figure 3, p. 117) was re-used to construct the fixed outer shape. The neutral levels of the attributes again consisted of 2-cm straight lines. Nose was now invariant, a vertical straight line placed centrally, identical to the former neutral Nose. All three Mouth forms were exactly as before -- one curve-up and one curve-down transformation along with the straight neutral level. However, the Mouth was repositioned: whereas before it was located half-way along the radius from the centre of the Nose to the circumference of the outer circle, now it was lowered to lie midway between the bottom of the Nose and the circle's circumference. (This was done simply because the lower Mouth "looked better" in the face configuration produced with the new Eyes.)

As to Eyes, in place of curves the two critical transformations now consisted of pairs of diagonal straight lines, each 2 cm long and set at an angle of $45^{0}/135^{0}$ from the (neutral) horizontal (see faces 3 and 4 in Figure 6, next page). In one Eyes pair (face 3) the two oppositely-inclined diagonals were widest apart at their upper ends and sloped down and in towards the top of the Nose; in the other (face 4), the two diagonals inclined towards each other at their upper ends and sloped down and outwards. All these diagonals projected equally above and below the location of the neutral horizontal Eyes bar. Regardless of variant, the two Eyes in each face were always sited 1 cm apart at their nearest points. (This meant that in the neutral level, while retaining the same straight-line form as in the second stimulus set, the two Eyes now lay slightly closer together on the horizontal plane than previously;

Figure 6

Third stimulus set -- used in Experiment 8.



however, their vertical distance above the Nose was as before, i.e., half-way along the radius from the centre of the Nose to the circumference of the outer circle.)*

The complete stimulus set contained five faces, illustrated in reduced size in Figure 6 above and reproduced in their original scale as well in Appendix 1, Figures 5 and 6. They comprised one critical Mouth pair, faces 1 and 2, equivalent to pair 12 of the second set; and one critical Eyes pair, faces 3 and 4, intended as affective counterparts to the former, non-affective pair 34. Since Nose was now irrelevant, the third set contained no equivalent to the previous Nose-differing pair 56. In its place was a single stimulus, face 5, neutral in all respects. All stimuli were mounted on the same (unframed) square, white cards as the second set.

Differentiation of the M-critical pair was expected to be associated with such terms as <u>happy</u>/<u>smiling</u> and <u>sad</u>, as before. Adults reacted to the E-critical pair with terms like <u>angry</u>/<u>frowning</u>/<u>scowling</u> for face 3 and <u>sad</u>/<u>anxious</u>/<u>worried</u> for face 4. (Adults universally agreed that they clearly conveyed something emotional and that the particular emotion expressed clearly differed between the two faces.) Pilot children, asked to describe faces 3 and 4, were able fairly readily to distinguish them as, e.g., <u>angry</u> versus <u>sad</u>.

---

* These Eyes forms were arived at partly by deliberate design, and partly by trial and error and the opinions of adults and children who were shown various versions. The aim was to keep Eyes as compatible as possible with Mouth in terms of physical structure and objective discriminability, while allowing the Eyes variants each to be associated with distinct affective connotations that could be readily appreciated and verbally labelled by young children.

The neutral face 5 was added for interest, to see whether it would be confused with or distinguished from the others in the set. This curiosity was aroused by the justificatory responses to the second stimulus set, in which the neutral Mouth was called either happy or sad, depending on what it was paired with. Most often, it was contrast that was captured: it was called sad in comparison with the curved-up (happy) Mouth and happy against the curved-down (sad) one. On occasion, however, the term indicated sameness; for illustration of such confusion of Mouths as both happy or both sad, see Appendix 3.2, Examples 21 and 23 (both produced by the same subject).

Stimulus presentation conditions

A total of 15 pairs can be generated from the five stimuli of the third set: five completely identical pairs, three differing in just Mouth (pairs 12, 15, 25), three differing in just Eyes (34, 35, 45) and four differing in both features (13, 14, 23, 24). The pair-comparison test session comprised all 15 possible trials.

Order of trials. Response sheets were drawn up listing the 15 trials in a different, randomly-determined order for each subject, with the proviso that pair 12 (M-critical) appeared on the very first trial in half the lists, and pair 34 (E-critical) appeared first in the other half. (Following the initial critical trial, the second critical pair could appear on any one of the next 14 trials.) This order-of-trials condition was added to permit a check on whether the subjects developed any "set" to keep on attending more to one attribute than to the other as a function of whichever had been the distinguishing feature on the first trial. The reason for considering this possibility was that the random encouragement of responding could reinforce neglect of a particular attribute if the child was praised early on for confusing it. While casual inspection of the data from the previous experiments had revealed no such tendency, it seemed worthwhile to investigate it systematically.

Stimulus alignment. In the second stimulus set, the two opposite curved forms of each attribute constituted mirror-image reversals of each other. Mirror images still obtain with the third set. Since faces 1 and 2 have simply been carried over almost intact from before, they again contain an up-down reversal of Mouth. The two new pairs of Eyes, in faces 3 and 4, are also up-down reversals of each other. To counter any mirror-image confusion effects, then, the Left-Right (L-R) and Up-Down (U-D) conditions of stimulus alignment were retained. Stimulus location within pairs -- on the left or right/top or bottom -- was randomised over trials and over subjects.

For practice, a selection was drawn from the practice stimuli used in the previous experiments.

Subjects

Thirty-two children were tested. Fifteen came from a community-centre playgroup in Oxford, six from a playgroup within the University of St. Andrews, and the remainder from another St. Andrews playgroup. Between them, they covered a wide range of home circumstances.

These subjects were assigned to subgroups to receive different stimulus presentation conditions, as follows. Half of the children viewed all pairs in L-R alignment, where the two stimuli were always displayed side by side. The other half received only U-D alignment, the two stimuli displayed one above the other. The ages of the 16 subjects in the L-R Group spanned from 3;2 to 5;2 years, with a mean of 4;1. The 16 in the U-D Group were aged from 3;1 to 5;2, mean 4;1 years. Each group was further subdivided: half to receive the M-critical pair 12 on the first trial, and half the E-critical pair 34 first. Age, sex and source of subjects were counterbalanced as far as possible across the various subgroups.

Procedure

In the university playgroup, stimuli were laid out on the floor in a quiet area of the room, and the child and Experimenter kneeled side by side on the floor to view them. In the other two centres, stimuli were displayed on a table set up in a corner of the room, with child and Experimenter seated side by side on chairs. Subjects were invited one at a time to look at some pictures.

The usual pair-comparison method was followed. On each trial, the appropriate pair was turned over and placed in front of the subject, who was asked, "Are these two (pictures) the same or not the same?". Verbal encouragement was given, but no feedback as to (in)correctness of judgments. Every child was asked to justify his answer ("How are they (not) the same?") to both critical pairs, 12 and 34, and to at least one of the five identical pairs. A reason was requested on an arbitrary selection (differing across children) of the other unidentical pairs. Altogether, each child was asked for a reason on about a third to a half of his trials, although children who were reluctant to answer were not pressurised into doing so. Additionally, a number of children spontaneously offered justifications when none had been requested. Preliminary practice was provided using the "non-face" stimuli, until the Experimenter was satisfied that the child understood the task and could discern all the relevant aspects of the practice figures. Four or five practice trials normally sufficed. The entire session took about 10 to 15 minutes per subject.

Prediction

Experiment 8 set out to explore the likelihood that error-making children might be assigning weights among the various face attributes on the basis of their affective significance, and taking (dis)similarity of expressed affect as the criterion in their pair-comparison judgments. This explanation for the previously-observed patterns of correct Mouth judgments and Eyes confusions was preferred to the criterial-attribute view that Mouth is universally treated as the more important feature in this kind of context.

To test this theoretical position, stimuli were provided in which variations in affective expression were carried by Mouth in some cases and by Eyes in others. According to the affect hypothesis, this should now enable Eyes variants to be discriminated as well as Mouth. The main prediction was therefore that the E-critical pair 34 would receive as many correct "not the same" judgments as the M-critical pair 12. The alternative hypothesis was that the E-critical pair would continue to receive fewer correct judgments than the M-critical pair. The latter outcome would be in line with the view that Mouth is criterial in itself, or might result if the attempt to capture affect in the structure of the Eyes were unsuccessful.

Of subsidiary interest were how the children would react to the neutral face 5, and how they would handle pairs differing by two features (i.e., where two sorts of affective messages were being conveyed), though no precise expectations were formulated on these questions.

RESULTS

Only twice in the entire experiment were actually-identical stimuli wrongly called not the same. Table 40 summarises the data for unidentical stimuli, showing the number of correct "not the same" judgments given to each of the 10 pairs, both overall and under the various subconditions separately.

Stimuli differing by two attributes were called the same 14 times out of a total of 128 opportunities for confusion here. While this represents a fairly low double-confusion rate (10.9%), it is slightly higher than those observed with the second stimulus set (5.8% in Experiment 4 with unfamiliar stimuli, and 2.4% -- only one error -- in Experiment 6 following familiarisation with the stimuli). This point will be returned to when the children's justifications for their decisions are considered.

But the third stimulus set was by no means associated in general with inferior pair-comparison performance. On the contrary, 14 out of the 32 subjects in Experiment 8 (44%) gave a correct judgment on every trial. Compare this with the 15% entirely correct in Experiment 4. Even among children familiar with the stimuli (Experiment 6), only 43% produced entirely correct sessions. Notice, further, that the average age of the Experiment 8 subjects, i.e., just turned four years old, was about six months younger than in the previous experiments. A younger sample was taken here since pilot tests showed that rising-five-year-olds produced "too many" all-correct sessions -- it was errors that were the primary interest. The success of even three-year-olds in distinguishing both Eyes and Mouth with the third stimulus set was remarkable. Correctness was not related to age or subgroup of subjects.

TABLE 40 (Experiment 8)

Number of correct "not the same" judgments for the ten non-identical pairs, by stimulus alignment and order-of-trials conditions.

SUBJECT GROUPS: Stimulus alignment/distinguishing feature on first trial	STIMULUS PAIR AND DISTINGUISHING FEATURE(S): critical pairs		non-critical pairs: 1-attribute differing				non-critical pairs: 2-attributes differing				total correct	(total correct
	M	E	M		E		M & E					
	12	34	15	25	35	45	13	14	23	24	actual	possible)
L-R: M first (n = 8)	8	7	6	6	8	7	8	8	8	7	73	(80)
L-R: E first (n = 8)	6	6	7	5	7	7	7	8	8	7	68	(80)
U-D: M first (n = 8)	7	7	7	7	7	7	7	7	6	5	67	(80)
U-D: E first (n = 8)	6	6	5	5	6	6	7	7	7	7	62	(80)
All L-R Group (n=16)	14	13	13	11	15	14	15	16	16	14	141	(160)
All U-D Group (n=16)	13	13	12	12	13	13	14	14	13	12	129	(160)
All M-first Group (n=16)	15	14	13	13	15	14	15	15	14	12	140	(160)
All E-first Group (n=16)	12	12	12	10	13	13	14	15	15	14	130	(160)
All subjects (N = 32)	27	26	25	23	28	27	29	30	29	26	270	(320)
% correct overall	84%	81%	78%	72%	88%	84%	91%	94%	91%	81%	84%	

The principal question in Experiment 8 concerned the proportions of discriminations of Eyes relative to Mouth on trials when only one of the attributes differed, the two critical stimulus pairs being 12 (M-distinct) and 34 (E-distinct). Was Mouth still dominant in correct judgments here? Of the 32 subjects, 24 succeeded in distinguishing <u>both</u> critical pairs, and three confused both. A further two children confused Mouth but not Eyes while three confused Eyes but not Mouth. Clearly, the Eyes and Mouth scores -- 27 correct on pair 12 and 26 on pair 34 -- do not differ significantly. The only subcondition to produce any difference between Mouth and Eyes scores was the L-R-alignment + pair-12-first permutation, where all eight subjects were correct on Mouth but one made an error on Eyes. The null hypothesis therefore stands unrejected: the critical Mouth-distinct pair and the critical Eyes-distinct pair were not associated with different proportions of correct judgments.

Responses on the <u>non</u>-critical trials involving a single-attribute difference (pairs 15, 25 for Mouth and pairs 35, 45 for Eyes) were examined by the Binomial test* (one-tailed, since the alternative hypothesis was directional, predicting fewer E- than M-discriminations). For all subjects combined, $p = 0.113$ ($N = 11$, $x = 3$); p-values associated with responses within the separate subconditions all exceeded that level. A further test was conducted between just pairs 25 (M-distinct) and 35 (E-distinct), the location of the greatest discrepancy between Eyes and Mouth scores: for all subjects, $p = 0.090$ ($N = 8$, $x = 2$). Although it

----

* Cf. the method of the McNemar test for the significance of changes; see Siegel (1956). A subject was classed as "changing from + to -" if he judged more of the M-pairs than of the E-pairs correctly, i.e., if he scored 2 (or, 1) correct on Mouth and 1 (or, zero) correct on Eyes. Conversely, he was classed as changing from - to + if he was correct on fewer of the M-pairs than of the E-pairs.

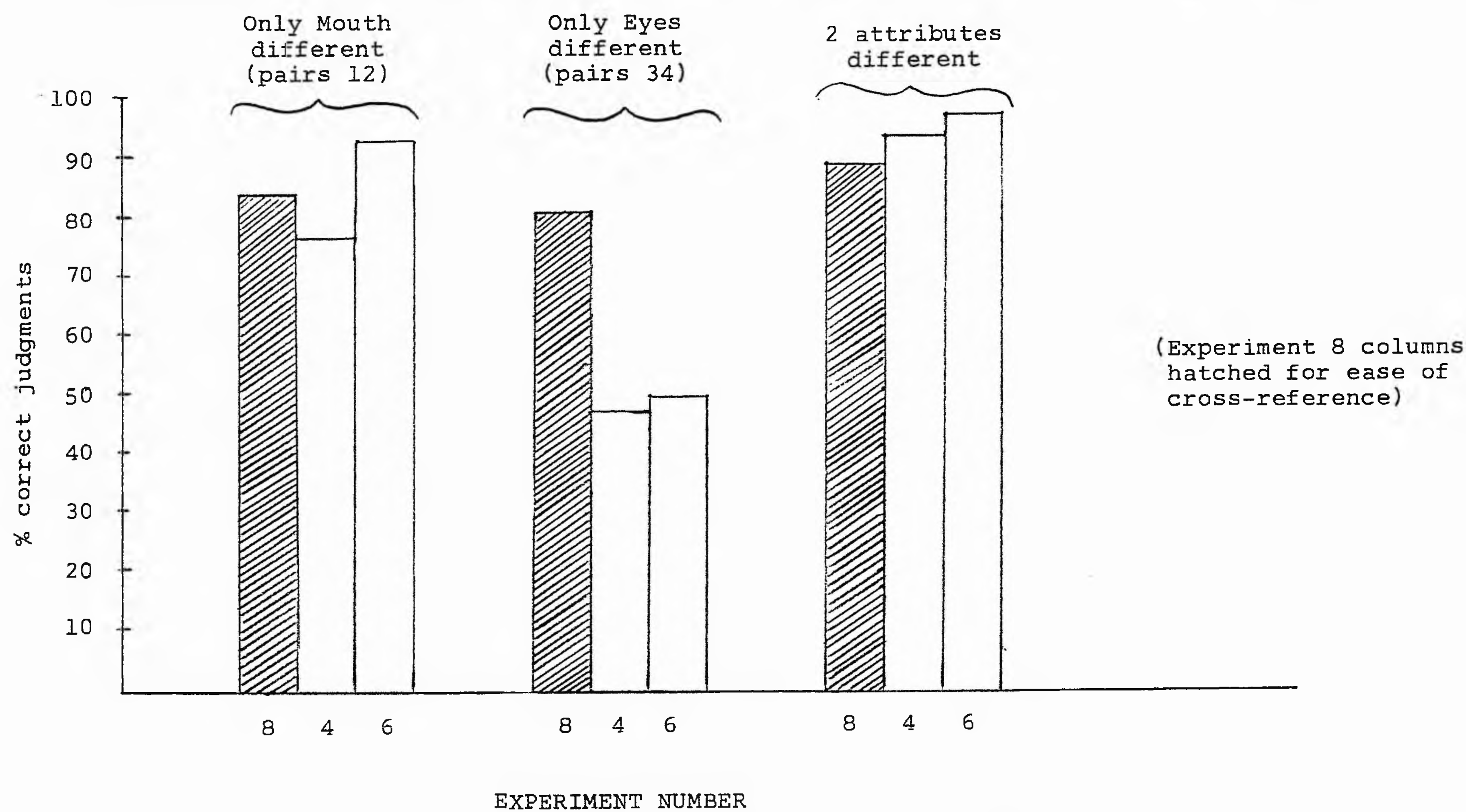
was not statistically significant, notice that the tendency lies in the opposite direction to that expected on the alternative hypothesis: these Eyes-distinct pairs were distinguished slightly more often than were the corresponding Mouth-distinct pairs (86% as opposed to 74% correct, overall). This was surprising after the results of the previous experiments, especially since all of these pairs had in common the inclusion of the neutral face 5, whose straight-line features one would have expected to be about equally physically discriminable both from the curved Mouth variants and from the diagonal Eyes variants. Either this assumption of equal objective discriminability was false, or the Eyes differences were more subjectively salient than the Mouth differences.

As a final check, a Cochran Q test was performed across all ten unidentical face pairs. The highest values of Q obtained were: for all subjects combined, 12.86; for that half of the sample given L-R stimulus alignment, 13.73; and for that half given the E-critical pair first, 11.93 (df = 9 throughout). All fail to reach significance. Frequency of correct judgments did not vary among the face pairs, whether the two stimuli differed by Mouth, by Eyes or by both attributes.

As far as the pair-comparison judgments themselves were concerned, then, the manipulation of the attribute variants in Experiment 8 appears to have been successful and the prediction based on the affect hypothesis was strongly confirmed. Performance was uniformly high across all stimulus pairs, regardless of which attribute had to be discriminated. The children were certainly attending to Eyes as much as to Mouth. The level of success attained here with Eyes is striking not only in relation to its parity with Mouth in the present experiment, but also in relation to the large proportions of Eyes confusions observed previously. Figure 7

Figure 7

Percent correct "not the same" judgments of attributes: comparison across Experiments 8 (E and M affective, stimuli unfamiliar), 4 (M affective, stimuli unfamiliar) and 6 (M affective, stimuli familiar) (all stimulus conditions combined)



highlights this, graphically comparing the overall percentages of correct judgments of Mouth and Eyes across Experiments 8, 4 and 6.

The stimulus-alignment and order-of-trials conditions need little comment since neither had much effect. Although performance was slightly worse in general (not significantly so) with U-D alignment, U-D subjects were not so far behind L-R subjects in Experiment 8 as they had been previously, despite the continued possible interference from the up-down mirror images of the attributes' variants. Similarly, subjects given the E-critical pair first fared generally slightly worse than those given the M-critical pair first. But there was no evidence whatsoever of an attentional bias as a function of the distinctive attribute first presented; on the contrary, <u>both</u> of these subgroups of subjects confused marginally <u>fewer</u> Eyes-distinct pairs than Mouth-distinct pairs overall.

<u>Justificatory responses</u>

From the pair-comparison judgments themselves, we cannot tell whether or not the high level of performance on Eyes relative to Mouth was related to the ascription of affective significance to the features. Is there any evidence from the children's justificatory responses? The justifications were anlaysed in the same manner as before (Part 4). Samples of these responses are provided in Appendix 3, Section 3. Only two children (four-year-olds) failed to answer at all under this section, and the response rate was high among the others -- almost 50% of all trials received a justification.

<u>Order of reference to attributes</u>. The children produced a total of 47 multiple responses, referring to both variable attributes and sometimes including the invariable Nose as well. Not surprisingly, the four face pairs that differed in both Eyes and Mouth attracted over half the

double references. Neither attribute was favoured over the other: Eyes were indicated first 22 times and Mouth was first 25 times. There was no obvious relationship between order of mention and the face pair referred to.

Number of references to each attribute. Again, neither attribute predominated. Mouth received a total of 117 references, Eyes a total of 121.* In single responses, which attribute was mentioned did depend on the particular face pair referred to: Mouth was what was indicated when that was the sole distinguishing feature (pairs 12, 15, 25) while pairs differing by Eyes (34, 35, 45) received references to Eyes. This is evident in Table 41.

Type of reference. References were classified according to the system established in Part 4 (pp. 181-185 and 204-205). Since a total of only three Point-level references occurred altogether, these are here subsumed under the Point/Name heading. The Feature-level and Point/Name categories were otherwise as before. The Face-level category, however, was split into references pertaining to an Affective state or action and those pertaining to some Other, non-affective, state or action. (For examples of the latter type, see Appendix 3.3, Examples 7, 13, 32.) The particular attribute with which a Face-level reference was associated is no longer specified, since pairs differing by a single feature were characterised in terms of that feature alone (so the referent in these cases is obvious), while for pairs differing in both attributes it was

---

* Additionally, there were several Face-level references that could not be attached to a single attribute specifically. These fell to pairs differing in both attributes or including face 5 (neutral in both attributes), where it was often unclear what was giving the face its "cross" or "happy" (etc.) character unless the child pointed to a feature.

often impossible to tell what the characterisation was founded on -- Eyes, Mouth, or the global configuration of both.

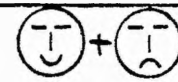
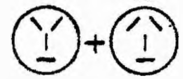
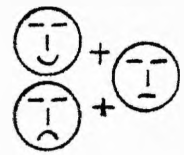
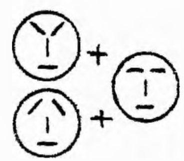
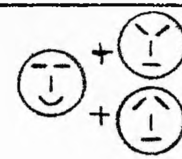
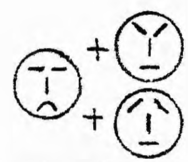
Table 41 shows how the references were distributed by type over the ten unidentical face pairs. Since frequencies were rather low, raw scores are presented, although the number of Affective Face-level references is also given as a percentage of all responses for each row to allow comparison across face pairs and with past experiments. For simplicity, face pairs have been grouped together on the basis of their kinds of distinguishing features.

To dispense with a minor point first, notice that compared with the previous experiments there were now relatively fewer Point/Name references and more Feature-level distinctions. Nor was pointing a resort of only the younger members of the sample. The features of the third stimulus set seemed generally easier than those of the second set for the children to describe adequately in words.

The main concern, however, is the affect-related responses. It is immediately obvious from Table 41 that pairs distinguishable by Eyes alone (34, 35, 45), while not reaching the same proportions of affective responses as the Mouth-distinct pairs attracted, nevertheless received Affective Face-level references about a fifth of the time and were successfully differentiated on that basis. This is a substantially higher percentage than was previously attained in the Face-level category with the second stimulus set. Before, only one or two Eyes references in an entire experiment were Face-level (considering only the pair-comparison tasks and not the familiarisation-period descriptions of Experiment 6). It is the figures in the present <u>Other (non-affective) Face-level</u> category that in fact correspond with the earlier findings, in terms of both

TABLE 41 (Experiment 8)

Number of justificatory references of each type according to face pair being judged

STIMULUS PAIR(S) AND BASIS OF DISTINCTION				TYPE OF REFERENCE						
				Affective Face-Level	Other Face-Level	Feature-Level M	Feature-Level E	Point/ Name M	Point/ Name E	Total no. Refs.
critical, 1 feature different	M	1+2		13 (45%)	0	6	1	7†	2	29
	E	3+4		5 (18%)	1	1	11	0	10	28
non-critical, 2 features different	M	1+5, 2+5		12† (30%)	0	8	1†	15†*	4	40
	E	3+5, 4+5		7† (21%)	2	2	8	4†*	11*	34
both M & E different	pos.+neg. affect ?	1+3, 1+4		16 (31%)	1	3	10	9	12†*	51
	both neg. affect ?	2+3, 2+4		12† (27%)	2	6	8	8	8	44
Percentage of each type of reference overall including identical-pair trials				30%	2%	10%	16%	20%	21%	262

Key:

† Up to three of the references in this cell were associated with pair-comparison confusion errors, although the references themselves were not usually erroneous.

* Up to two of the references in this cell were erroneous, though not necessarily on wrongly-judged trials.

frequency and content -- the previous Face-level references to Eyes did not concern affect. In both cases, phrases like "he's closed his eyes" and "he's pulling his eyes down like this" occurred.

Affective Face-level responses were even given to identical pairs, where there was no contrast to highlight the affective significance of the face or feature: pairs 11 and 22 together received 13 justifications of which seven were affective; pairs 33 and 44 received five affective descriptions out of 13 responses. Even the neutral face 5 when paired with itself received two affective descriptions, out of ten responses (presumably because a "set to respond affectively" had been established with these stimuli?).

Affective descriptions given on trials in which a pair-comparison confusion error was made are interesting, for although a wrong "same" judgment was given the descriptions themselves could not be called erroneous. The presence of acceptable descriptions associated with confusion errors is marked in Table 41 by a superscript "†": they fell either to pairs that included the neutral face 5 or to pairs differing in two features. Instances of acceptable descriptions associated with such confusion errors are given in Appendix 3.3 (e.g., Examples 22, 34, 50, 55 -- an asterisk against the example in the lists indicates a wrong pair-comparison judgment).

In the case of the neutral face 5, we already noted the possibility that it might be taken as either happy or sad (p. 242). Most often, it was called by whichever term would contrast it with its partner in the pair; but sometimes it was ascribed the same affective character as its partner. The list in the Appendix provides a number of illustrations of the switching of face 5 between happy and sad, with the occasional resulting confusion error. For instance, Examples 16, 23, 28 and 34 were all produced by the same subject, who called face 5 sad against face 1, happy against face 2, happy against face 3 (almost a wrong "same" judgment, but self-corrected!) and sad against face 4 (with a wrong "same" judgment).

With respect to the pairs differing in both Mouth and Eyes, we have already seen that a slightly greater proportion of confusion errors occurred than would have been expected on the basis of performance on two-attribute differences with the second stimulus set (p. 246). Although at the outset of the experiment expectations had been left open as to how the children would react when the stimulus pair incorporated two sorts of affective signals (i.e., differed in both Eyes and Mouth), a speculation was entertained, which has been entered in Table 41. There, it is suggested that pairs combining face 1 with either face 3 or face 4 might be distinguishable on the grounds that face 1 represents positive affect (happy, smiling) while faces 3 and 4 are interpretable in negative affective terms (cross, sad); while it is speculated that pairs combining face 2 with either face 3 or face 4 might be indistinguishable since all three faces could be taken to signify negative affect (all cross, sad). In reality, face 3 turned out typically to be called happy to distinguish it from the negative (sad/cross) face 2. But confusions of the negative-negative type did occur with pair 24, where each was called cross, or each sad. This tendency is reflected in Table 41 in the lower percentage of correct "not the same" judgments of pair 24 relative to the other Mouth-plus-Eyes-distinct pairs.

These confusion errors involving the neutral face 5 and pair 24 demonstrate the strength of the influence of affect for some children.*

---

* The discussion here pertains to acceptable affective descriptions of the faces in conjunction with wrong "same" pair-comparison judgments. Erroneous references on unidentical pairs were extremely rare, and amounted

In these cases, apparently, the children's justificatory comments were not simply post-hoc reasons, concocted after the event of pair-comparison decision. They had, rather, the character of reasonings, revealing something of the cognitive processes whereby the child arrived at his decision in the first place. It appears as though the stimuli of the third set were being not only talked about in affective terms, but perceptually-cognitively encoded in those terms, too. Thus, the results suggest, it was precisely because the faces were being interpreted in affective light that they were confused in certain instances and successfully differentiated in others.

Summary and conclusions

Experiment 8 followed the conventional pair-comparison task method employed in Experiments 3 to 6, but with a new set of schematic faces as stimuli. It was predicted at the outset of Experiment 8 that, in contrast with the previously-observed tendency for Eyes to be confused, Eyes would now be correctly discriminated as often as Mouth. This was very clearly confirmed. The level of pair-comparison performance was uniformly high on both Mouth-distinct and Eyes-distinct pairs. Even three-year-olds could perform entirely correctly. The Eyes variants were differentiated more than 80% of the time, where before the success rate on this attribute was consistently less than 50%, despite the slightly older

---

/continued/ to simply saying "The mouths/eyes are the same" (etc.) when the feature actually differed. The presence of such erroneous references is marked in Table 41 by a superscript "*" against the appropriate Point/Name column cells. Only the pairs including face 5 or differing by two features incurred such false statements. Given the occurrence of affect-based confusions on these face pairs, there is some doubt as to whether the Point/Name references really were erroneous. The children who said things like "They've got the same eyes" might have meant "Their eyes convey the same kind of affect" rather than "Their eyes have the identical physical form".

subject samples in the earlier experiments. This achievement with the third stimulus set was especially surprising in view of the physical structure of the Eyes variants here -- pairs of diagonal lines forming mirror-image reversals of each other. In other studies, as we have already seen (pp. 8ff), diagonal straight lines differing in orientation, and particularly in mirror-image reversal, have been found to be especially confusable for preschool children. Evidently, there was something about the face-like visual context in which the present diagonal lines were displayed that facilitated their differentiation.

The prediction that Eyes and Mouth differences would now be correctly judged equally often had been founded on the hypothesis that the affective significance of the features in the kinds of face stimuli employed in the present study might be important to the children in their sameness-judgment task. Over all subjects, justifications for pair-comparison decisions were produced on a total of 215 trials; 37%* of them consisted of a description of one or both stimuli of a pair in terms that characterised not physical identity/difference between attributes individually, but the global affective expression of the whole face.**

---

* This figure of 37% is higher than that quoted in Table 41 (30%), because in the latter case Feature-level and Point/Name responses that were multiple (i.e., that referred to both Mouth and Eyes against a single trial) were counted as two separate references. This brought the total on which the Table 41 percentage was calculated to 262 instead of 215. Face-level references, however, by their very nature, did not usually mention Mouth and Eyes separately, and were scored as only one reference even when they did. It is therefore a fairer comparison, for present purposes, if multiple Feature-level and Point/Name references to a single pair are also counted as only one response per trial.

** In this connection, it should be noted that the third stimulus set received very few justifications of the type "They're the same <u>but</u> the eyes are not the same" or "The eyes are not the same <u>but</u> the faces are the same because they're both happy" (etc.), which were common with the first and second stimulus sets (see pp. 212ff). This is probably partly because comparatively few confusion errors now occurred; but in those that did arise, the child usually focussed solely on what was <u>shared</u> in the faces (i.e., typically the part conveying sameness of affect).

Half of the subjects produced two or more Affective Face-level references, showing that they were attending to affect <u>at least</u> some of the time. (And this does not take into account the possibility that some of the Point/Name references -- "The eyes are not the same", etc. -- might have been founded on affective rather than physical grounds.) To demonstrate just how pervasive the influence of the affective character of the stimulus faces could be, records of the complete 15-trial sessions of two of the subjects have been added to Appendix 3.3 (p. xxiii).

More importantly, Affective Face-level references no longer fell solely to Mouth. Of those cases where it was clear which of the attributes the child was taking as carrying the affective expression, just over a third were associated with Eyes. Although Eyes were not yet on a par with Mouth in terms of proportion of Affective references,* then, it seems probable that there was a connection between the increase in correct Eyes discriminations and the fact that Eyes could now be distiguished on affective grounds.

The children's justificatory responses further suggested that the effect of the experimental manipulation of the affective significance of the stimulus features was <u>twofold</u>. While it likely contributed to the overall improvement in Eyes discriminations, paradoxically it is certain that it also contributed to some of the observed wrong "same" judgments. Especially telling in this regard were the false "same" judgments given to stimulus pairs that actually differed in both Eyes and Mouth. Such

---

* It is also possible that the attempt to lend affective significance to Eyes as well as Mouth might not have been entirely successful. From a child's point of view, the simple happy-sad (positive-negative) dichotomy afforded by the curve-up and curve-down Mouth forms might have been more salient than the relatively more subtle cross-sad (negative-negative) distinction between the diagonal Eyes variants.

double-attribute confusions were rare in the earlier experiments, and in the present experiment even the number of confusion errors when there was only a single attribute's difference was small. On those grounds, one would have expected double-attribute confusions now to have been minimal indeed. That they were not virtually absent is perhaps one of the most interesting findings of this entire series of experiments -- at least when the children's justifications for their decisions are also taken into consideration. Faces 2 and 4, for example, were called the same a fifth of the time, being described as <u>both cross</u> or <u>both sad</u>, clear indication of <u>why</u> the two stimuli were being confused.

At the start of the final part of this study, two explanations for the patterns of results obtained in the preceding experiments were considered: (1) that Mouth is a universally important (criterial) attribute such that it would consistently receive greater attention, and greater proportions of correct sameness-difference judgments, than other attributes of schematic faces; (2) that what was important here was affect, and that <u>any</u> feature carrying affective significance would be "criterial". The results of Experiment 8 argue convincingly against the first of these two possibilities: Eyes can be differentiated just as often as Mouth under certain conditions. That the variants of an attribute be differentiable on the basis of affective significance is one of those conditions is strongly suggested from the patterns of pair-comparison responses and the types of justifications given by the young subjects in this last experiment.

OVERVIEW AND CONCLUDING REMARKS

The present research programme began with the assumption that young children make "confusion errors" when they are given tasks requiring them to judge sameness and difference among visual stimuli. The occurrence of confusion errors is widely documented elsewhere. In experimental tasks such as matching-from-sample (with the instruction, e.g., "Find one the same as this one") they often choose a mismatch instead of the perfect match to the standard; or in pair-comparison ("Are these two the same or not?") they call items the same when they exhibit points of difference. Confusions are also observed in tasks involving serial recognition, discrimination learning, and so on. For adults given these kinds of tasks, things are the same only if they are <u>identical</u> in all visual features. Why then do children call things the same when they are not physically identical? Do they not yet fully understand what is meant by "same"? Or does the source of their confusion errors lie rather in the perceptual-cognitive processes of visual inspection and comparison across stimuli?

One interesting fact to emerge from the literatures on children's sameness-difference performance is the <u>systematicity</u> of their confusion errors. They are fairly likely to occur with certain types of discriminanda, and are unlikely with other sorts of discriminanda. Thus, for instance, stimuli that differ only by orientation seem highly confusable, while distinct shapes are rarely confused. Such findings point to the operation of (at least) perceptual-cognitive factors in the production of errors.

This then formed the background to the present study. The initial brief was to carry out experimental manipulations of visual aspects of the stimuli in sameness-difference judgment tasks, in order to examine some of the effects of particular perceptual factors on the differential distribution of confusion errors among various stimulus attributes.

The starting point was provided by the seemingly contradictory results of John Taylor (1973) and Roger Wales (pers. comm, 1974). Those investigators both employed a stimulus set consisting of eight schematic faces, originally designed by Tversky and Krantz (1969). In both cases, the faces were constructed from three attributes each taking two levels: a horizontal or vertical ellipse for Shape of head; black or blank-white circles for Eyes; and a straight line or an upward curve for Mouth (see Appendix 1, Figure 1). It seemed possible that the discrepancy between Taylor's finding of correct Shape matches with internal feature confusions and Wales's report of correct internal feature matches with Shape confusions might reside in whether or not the stimulus cards incorporated a square frame drawn around the face. Specifically, the proposal was that the presence of an external visual frame (as in Taylor's stimuli) might be associated with correct matching of the outermost feature, Shape, while in the absence of such a frame (as in Wales's stimuli) Shape might be more confusable than the internal features Eyes and Mouth.

The three experiments in the first part of the present study set out to test this explanation. Children from preschools and early primary school classes were presented with the Taylor and Wales schematic faces, either Framed or Unframed. Their patterns of confusion errors were investigated under three task conditions: matching-from-sample (goal, to find from an array the only perfect match to the standard); choice-from-triads (choose from three unidentical stimuli the two that look most alike); and pair-comparison (decide whether two stimuli are the same or not).

The results clearly showed that presence/absence of a surrounding frame was not the sole determinant of which attributes were correctly matched and which confused. There was an interaction between the visual Framed/Unframed condition and the type of task used to elicit sameness decisions from the children. When all of the stimuli in the set were on view simultaneously in a large array, as in matching-from-sample, Shape was correctly matched and Eyes were confused regardless of whether a frame was provided or not. In contrast, when only two stimuli were presented at a time for pair-comparison, the tendency was more towards internal feature matching, especially with Unframed stimuli. The "intermediate" triads task revealed both effects clearly -- Shape matching with Framed and Mouth matching with Unframed stimuli. The discrepancy between Taylor's and Wales's findings is thus explicable only if both the methodological and frame conditions are taken into account, in combination. For, Taylor's Shape-matching subjects received a matching-from-sample task with the Framed stimuli, while Wales's internal feature matches were elicited under pair-comparison with Unframed faces.

Exactly why this interaction of frame and task variables should have produced the observed effects on the distribution of confusion errors among the stimulus attributes remained somewhat puzzling. One might conjecture that the large array of the matching-from-sample method, in which the eight stimuli plus standard were displayed adjacently in a 3 x 3 square arrangement, perhaps made the two distinct Shape levels highly noticeable, even without the frame, because of their proximity in numbers. In the smaller choice array of the triads task, Shape was then presumably less salient and the originally-predicted frame effect came through. That is, among Unframed stimuli choices were based on sameness of Mouth; but the presence of an external frame apparently again highlighted the outer Shapes of the faces.

The above proposals rest on the difference in stimulus array between the matching-from-sample and triads methods. In the pair-comparison task, resembling triads in its small array, however, the Framed versus Unframed effect was less marked. Although Unframed stimuli elicited the predicted internal feature matching, the Framed condition did not produce the expected contrasting superiority of Shape -- Mouth was also correctly matched. One must now appeal to the effect of the task requirements as well as of the visual array. Pair-comparison differs from both of the preceding methods in requiring "passive" visual scanning of the figures and verbal Yes/No answer, rather than active choice of one or two stimuli from multiple options. Yet precisely how such factors would operate is not at all clear, and work remains to be done before the complexities can be disentangled. Beyond noting some of the problems to be dealt with in the future, this line of inquiry was not extended here. The findings from the first part of the study do suggest, however, that the literatures on children's errors in sameness judgments and discrimination learning tasks (etc.) need to be interpreted with caution, in view of the diverse methodologies -- as well as the variety of discriminanda -- employed by different investigators.

The second part of the study examined the role of differential visual saliencies among the stimulus attributes from a new angle. It was possible that the bias of confusion errors towards one rather than another of the attributes in the first experiments might partly have resulted from the widely differing physical structures of the original three attributes.

Throughout Experiments 4 to 6, the task was held constant (pair-comparison). A second set of schematic faces was constructed (see

Appendix 1, Figure 3) in which Shape was invariant and the relevant attributes were all internal: Mouth, Eyes and Nose. The aim was to determine whether confusion errors were due to failure to perceive an attribute or failure to detect its distinct levels, or instead due to deliberate selection of one attribute and rejection of another as the basis for sameness decisions. This question was examined by equating physical structure across the attributes as far as possible. Each attribute had the same three variants: one neutral straight line, and two curved lines oriented in opposite directions.

Through this section, Mouth dominated the correct judgments, being successfully discriminated as much as twice as often as were Eyes or Nose. This pattern held even when the stimuli were presented upside-down, or when the subjects received "pretraining" to familiarise them with the discriminanda and to ensure that they could discern the variants of each attribute. Visual factors such as mirror-image reversal effects and bias due to location of gaze and direction of scanning were eliminated as contributors. Also discounted was the "objective physical saliency hypothesis", by which Eyes would have been expected to be more noticeable than Mouth, there being <u>two</u> Eyes and only one Mouth in each stimulus.

The "<u>subjective</u> weighting (saliency) hypothesis" -- that the subjects were deliberately treating Mouth as a more important determinant of sameness-difference among these schematic faces, and were deliberately treating the other attributes as less relevant -- stood unrejected. This hypothesis, moreover, received active support from the children's verbal comments on the stimuli and justifications for their judgments, examined in Part 4. These showed that often children who confused Eyes and Nose were able to discern differences there, but were disregarding them in their sameness judgments. In this regard, comments from the subjects such as, "They're the same but the noses are going one that way and one that [the other] way" or, "The eyes are not the same but the faces are the same because they're both happy" were quite typical. As well as confirming that the child is intentionally treating Eyes and Nose as relatively unimportant, this sort of response refutes the claim that young children do not fully understand the word "same" itself. Rather, it demonstrates a quite mature appreciation of the <u>various</u> senses in which "same" can be used in the adult language.

As with the pair-comparison judgments proper, then, the children's justificatory responses pointed to the weight they were assigning to Mouth over the other attributes. Where the other features were indicated by gesture or name or at most by a description of their purely physical appearance, the Mouth was often described in terms that took it to characterise the whole face by action or mood (smiling, sad, etc.). Indeed, "mouth" and "face" were frequently treated as synonymous, sometimes to strange effect (as in "It's got a straight face", meaning the <u>Mouth</u> was a straight line). This never happened with the other attributes, and it points once again to the special import of Mouth here. The terminology of the subjects' justifications suggested that what they were focussing on might be the <u>affective</u> significance that could be ascribed to the particular Mouth forms (but not to any of the other attributes) employed in the first two sets of stimulus faces.

The final experiment of the series was designed to decide between two competing possible reasons for the children's weighting of Mouth over the other internal features of the faces of the first two stimulus sets. One, based on the results of the present study, was that it was not sameness-difference of Mouth itself that was crucial, but sameness-difference of affective facial expression (which up to now had been conveyed by Mouth alone). The alternative explanation was that, at least in such elementary schematic drawings, Mouth is for young children a criterial attribute of faces in its own right.

To test this, a third set of stimulus faces was devised in which Nose as well as Shape was held constant, while Mouth and Eyes varied (see Appendix 1, Figure 5). The Mouth forms were retained from the second set, but now the Eyes variants were constructed in such a way that they too could be distinguished on affective grounds (as, e.g., angry versus sad). The task was again pair-comparison. The results clearly countered the view that Mouth is a criterial attribute, for Eyes were now correctly differentiated as often as Mouth. The children's references to affect in their justifications further supported the argument that their pair-comparison decisions relied on whether or not the faces were alike in type of mood or emotion expressed. Several of the (few) pair-comparison confusion errors that were made were clearly due directly to the ascription of affect. Thus, for example, faces were called the same because "they're both cross", even though the pair differed structurally in both Eyes and Mouth, Eyes conveying the "crossness" in one face and Mouth conveying it in the other. This was convincing evidence that judgments were being made independently of the precise physical forms of the stimulus features except insofar as they permitted particular kinds of affect to be ascribed to the faces.

Evidence in support of the idea that affect can be important in the perception, classification and discrimination of face stimuli exists in other sources. With regard to schematic representations of faces, Pick _et al._ (1972) found that children confused lines curving upwards in various degrees more often than they confused curved-down lines in a learning/recognition task, whether the curves appeared as mouths in faces or whether they were presented in isolation. They suggested that the inferior performance on curved-up lines might be due to the children's classing them together as all "smiling mouths" among which they did not differentiate. Interestingly, they reported instances of the children's referring to the curved-up lines as "smiles" even when they were presented out of the face context. This suggests that "smiling" may be a very fundamental concept for young children.

Similarly, Hess and Pick's (1974; see also pp. 108-110 above) finding of more pair-comparison confusions on mouth than on eyes could be attributed to interference from affective interpretation of the stimulus features. In that study, the mouth forms were all curved-up lines, which again appeared to be treated as indistinguishable because they all belonged to the same class of "positive-appearing mouths". The eyes, however, were constructed from ellipses of varying degrees. This meant that not only did the shape of the eyes vary quite widely across the variants according to their curvature, but also the space between the eyes varied -- which, as Hess and Pick noted, may alter the global character of the facial expression. The superiority of eyes over mouth in Hess and Pick's experiment thus fits neatly into the present framework.

Other authors who have found that children attend mostly to the upper, eyes part of the face when the stimuli are photographs of real faces have also attributed this bias to the importance of facial expression or similar social factors (e.g., Goldstein and Mackenberg, 1966). Langdell (1978) reported that, unlike various groups of control subjects, younger (nine-year-old) autistic children identified photographed faces not from the upper portion of the face but from the lower half. He claimed that autistic children do not see the social character of faces but attend only to their visual, physical aspects; therefore they focus less on the socially-informative eyes. Odom and Lemond (1974) found that kindergarten children, given the task of recalling the position of a photographed face in a large array of such stimuli, performed better when they used facial expression than when they used personal identity. And Ahrens (1954) found that babies, looking at real faces, began by focussing on the eyes but soon switched to the mouth. Then, after about eight months old, the infant switched again, now attending to global expression instead of to individual facial features, responding positively to a smiling face and negatively to a frowning one.

Young children may also give precedence to minor physical differences over larger (more objectively salient) ones when the former involve affective distinctions and the latter do not. This emerged from the investigations of Chalmers and her colleagues (1980 and pers. comm.)[5] into the relationships between children's apparent referential inadequacies (cf. Glucksberg et al., 1966; Flavell, 1975) and perceptual and contextual stimulus variables. Children of various ages were presented with schematic drawings of people, in description and matching tasks. Members within a stimulus set varied in height (tall-short), hair colour (black-white) and mouth (curved up-curved down). In objective physical terms, overall size and hair were "large" features while mouth was a comparatively tiny one, and the differences between the two levels of height and hair colour would be expected to be considerably more salient than the difference between the two levels of mouth. Size and colour variations, moreover, are generally agreed to be less confusable for young children than is the kind of up-down mirror-image reversal inherent in the two mouth variants. Yet the responses demonstrated that mouth was _highly_ salient for the younger subjects. For example, when the figures were presented individually for free description, the first of the three relevant attributes that four-year-olds mentioned was the mouth (although this was not so true for seven-year-olds). References to mouth, moreover, were not descriptions of the objective visual characteristics of the feature, but were couched in terms of whether the _person_ was _happy_ or _sad_. That is, in line with the findings of the present study, mouth references were of the global Affective Face-level type instead of describing the physical Feature-level.

Even for adults, affect may be a salient aspect of schematic faces. This was evident in the classic studies of Brunswik and his colleagues (e.g., Brunswik & Reiter, 1938). He was particularly interested in the "impression value" of the face features, and which physical variations were associated with differences of mood and character, etc. (For a summary of this work, see Brunswik, 1956, especially Chapter XV.) Despite the highly simple, schematised nature of the Brunswik faces, subjects could readily describe and rank them according to seven categories of global "impression characteristics": mood ("gay" versus "sad"), age, character, likeability, beauty, intelligence and energy. In an analysis of the extent to which subjects' ascribed characteristics depended on the actual physical features of the faces, Brunswik found that it was variation of _mouth_ that produced the widest range of responses, which were especially extreme in the "mood" category of impression. Brunswik also discussed (1956, pp. 111 & 113) a study by Samuels (1939) who found that in pair-comparison with Brunswik faces the highest percentages of correct discriminations (by adult subjects) were associated with distinctions of mood, along with beauty, out of all the seven impression characteristics.

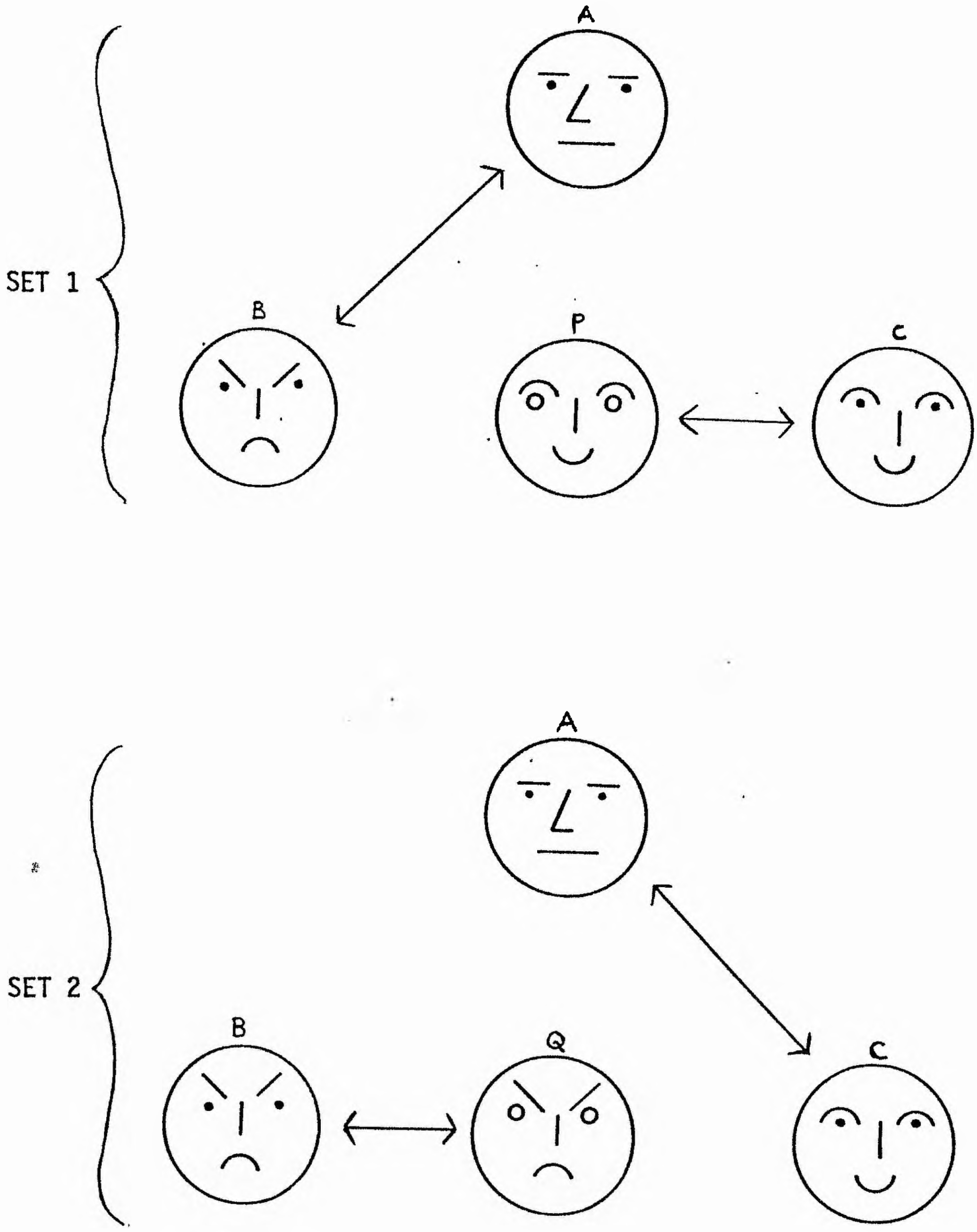
Within the domain of affect, some types of emotions are more salient for adults than others. The dimension happiness-sadness or happiness-worry seems to be especially important (e.g., Frijda, 1969). Bassili (1979) found happiness and surprise, indicated by mouth movements, to be the most readily identifiable expressions. This might explain why, in the final experiment of the present series, the children gave a greater proportion of Affective Face-level references to Mouth than to Eyes: Mouth captured the readily-recognisable happy/sad dichotomy, while the distinction between the affective expressions of the Eyes was more subtle.

The results of an experiment conducted by Tversky (1977, pp. 341-342) also comment on the fact that adults may take affect as the basis for classification and similarity judgments of schematic faces. His two stimulus sets are reproduced below (Figure 8). They differ in only one stimulus: the "happy" face P in Set 1 is replaced in Set 2 by the "cross" face Q. Instructed to sort the four stimuli into two pairs, adult subjects given Set 1 produced the pairs AB and PC. But subjects given Set 2 did not group A and B together; instead, they divided the set as AC and BQ. Further, when fresh subjects were asked to choose from the array B,P,C or B,Q,C the one that was most like the standard face A, face C was the clear choice from Set 2, but with Set 1 face B was favoured.

Notice that faces B and C each share only the eyes with the standard A. The subjects' responses must therefore have been founded on

Figure 8

Two sets of schematic faces presented to adult subjects for classification and matching-to-standard (Tversky, 1977, p. 341)



(Arrows have been added to indicate the most common partitioning of the stimuli in classification and the stimulus chosen as match to the standard face A in similarity judgment.)

something other than points of physical sameness and difference. Evidently, those given Set 1 were grouping the stimuli as (in Tversky's terminology!) "smiling faces" versus "nonsmiling faces", while Set 2 stimuli were treated as "frowning faces" versus "nonfrowning faces". That a change in only one member of the stimulus set could produce such a striking shift in adults' similarity choices supports the case for the importance of affect in the perceptual and cognitive encoding of faces. It also reveals the operation of a kind of "chain reaction": stimulus context determined the basis of classification, which in turn influenced the basis of individual similarity judgments. One could in fact say that there was no real shift in the subjects' decision-criterion across the two stimulus sets -- similarity of affective facial expression was used in both cases.

The emphasis placed here on Affective Face-level responding is not, of course, to be taken as saying that young children will inevitably match face stimuli according to global physiognomic similarity. In the present study, many of the Mouth/affect-matching subjects demonstrated their awareness of other features of the stimuli. It is likely that differing mouth forms (= differing affect) would be ignored if the visual saliency of other, non-affective attributes were increased or if there were a large enough number of other points of physical identity between the faces.*

---

* Cf. the demonstrations of the "dimensional additivity" effect by, e.g., Tversky and Krantz (1969), Gati and Tversky (1982): as the number of aspects of physical identity between stimuli is increased, but the number of differences remains constant, so the stimuli are more likely to be considered the same. It is as if the increase in actual (and hence, perceived) sameness brings about a reduction in the perceived difference even though the actual difference is no smaller. Cf. also Carey and Diamond's (1977) finding that young children have difficulty recognising a photographed face when extra-personal "disguise" features are added -- hat, spectacles, etc. While those paraphernalia did not hide the face, they appeared salient enough to "distract" the child away from the face itself.

The experiments of the present series have themselves shown that there are circumstances in which young children do not take shared affective expression as the criterion of sameness among the faces. Certain task and stimulus variables (finding a match to a given standard from a large array; presence of an external visual frame) were found to lead to higher proportions of correct matches of the outer head Shape than of the internal features, even when the confused Mouth forms could be interpreted as affectively different.

When the child did treat shared affective expression as a necessary criterion for sameness, it was often not his <u>only</u> criterion. When the stimuli contained three attributes -- Mouth, Eyes, and either Shape or Nose -- wrongly-matched face pairs tended to differ by just <u>one</u> attribute. That is, (where possible) an "affect-attending" child would typically decide on the basis of sameness of Mouth (affective) <u>plus</u> one of the non-affective attributes. Confusions of two attributes at once were especially infrequent with the second stimulus set. (The third stimulus set was exceptional in provoking some double-attribute errors, where similarity of global facial expression overrode physical differences.) The children did tend to take into account the <u>number</u> of points of physical identity as well as the <u>manner</u> in which the stimuli were alike, to the extent that those two criteria did not yield conflicting answers.

Setting aside the complications introduced by the manipulations of task and frame variables in the first three experiments, the present findings have implications that go beyond merely adding to the list of types of discriminanda that children are more, or less, likely to confuse. They offer fresh comment on what may and what may not underly these sorts of confusion errors. For one thing, it was verified that the children were assigning differential weights among the stimulus attributes on subjective rather than objective grounds, where in the past subjective weighting factors have often been overlooked or simply assumed in theory but rarely rigorously examined in practice (cf. Goodman, 1970; Taylor & Wales, 1970; Hale & Lipps, 1974; Gregson, 1975; Tversky, 1977). In the present instance, features that did not contribute to (variation in) facial expression were deliberately assigned low priority and their differences ignored even though they were detected. This in turn shows that the problem here was not simply a perceptual one of inability to discern features or failure to inspect stimuli fully enough.

It also suggests that Kemler and Smith's claim that young children operate in terms of <u>overall</u> similarity and only gradually come to differentiate individual stimulus properties should not be interpreted too simplistically. (This theoretical standpoint and the empirical research on which it is founded were reviewed in some detail in Part 1, pp. 25-28.) Smith and Kemler (1978) themselves note that certain dimensional axes may be primary and highly salient to very young children, despite their apparent tendency to perceive globally or integrally. The present results suggest that attending to "global" similarity may, paradoxically, entail focussing on just a part of the stimulus figure instead of the whole. One might, indeed, speculate that children's reputed failures on the side of perceiving and detecting -- their "inadequate" or unadultlike paths of visual inspection of stimuli, for example (e.g., Vurpillot, 1968; Vurpillot <u>et al</u>., 1971) -- may not stem from immaturities in perceptual functioning but may be the result of deliberate scanning strategies. If what the child is (intentionally) looking for is global likeness in a particular respect, he does not need to inspect every detail of the stimulus figure thoroughly.

This area of research also points to the danger of assuming tidy, structurally-specifiable and -measurable relationships between the physical properties of the stimuli and people's perceptions of them (cf. the psychometric approach -- see pp. 28-30 and the introductory sections of Part 3). John Taylor described the eight schematic faces he used in his third experiment (1973) -- those forming the stimulus set for the first part of the present programme (see Appendix 1, Figure 1) -- as composed of orthogonal elements that vary discretely. Even though that may be the case structurally speaking (which is debatable), in terms of people's actual behavioural interactions with those stimulus faces they may be treated as if composed not of separable but of integral attributes, such that a change in the physical characteristics of a single feature may serve to effect a change in the way the entire stimulus is perceived.

What about the language side of the problem, the child's understanding of the word "same"? In adult speech, this word has multiple usages (which is appropriate in a given instance normally being clear from the linguistic and non-linguistic context). It can refer to a single entity that re-appears in space/time; it can indicate that two or more entities are physically identical in all respects; or it can indicate that two or more entities, while perhaps not physically identical, belong to the same class with respect to some principle, i.e., share some quality, function, etc., that defines the class. For the sake of brevity, let us call the third sense qualitative sameness; the second, by contrast, could then be regarded as involving quantitative sameness.*

---

* This distinction is akin to that made by Gregson (1975, Ch. I), who took "quantitative similarity" to be determined by counts of numbers of shared physical features, and "qualitative similarity" as involving those of the "potential similarities" among stimuli that are actually used by people in

In laboratory tasks of the kind employed here, adults universally adopt the second, quantitative interpretation, judging items to be more and less the same according to the number of shared physical features. The many children who exhibited consistent error patterns in the present study, however, seemed to be deciding primarily by qualitative sameness and only secondarily by quantitative sameness. Since their verbal comments indicated that they were able themselves to apply the word "same" appropriately in each of its senses, they must have been deliberately choosing to take the Experimenter's test question to refer to "same kind of thing".

This conclusion is corroborated by the work of Karmiloff-Smith (1977), who found that French-speaking children also interpreted "same" (même) to mean "same kind", but this time in a situation where the adult interpretation was the first of the three senses distinguished above, i.e., when it was intended as "the very same one". For example, when asked to act out what happened when "The boy pushed a cow and then the girl pushed the same cow", children below age five pushed not one cow twice, but two separate cows. Interestingly, while three-year-olds limited their responses to two identical cows, four-year-olds often chose two unidentical cows (e.g., cows differing by colour). But unlike younger children, four-year-olds did alter their responses to the "very same one" if the Experimenter added some further clarification, such as "... the same cow as the boy just pushed". Further, when given two toy ducks differing only in colour, some of these children answered No to the

---

/continued/ a given situation. It is not to be confused with the terminology of Lian (1981, pp. 43-49), who called the first of the three domains "numerical identity" and apparently confounded the second and third under "qualitative identity".

the question "Is it the same?" but Yes to "Is it the same duck?".* This again illustrates that they did appreciate that "même", like "same", has diverse senses, even if they were as yet unable to match sense and context appropriately. By age five, Karmiloff-Smith found, children could correctly perform the task requiring them to push one-and-the-same cow twice. In their spontaneous comments, moreover, they differentiated the two senses of "same" ("very same one" versus "same kind of thing") by using ungrammatical periphrastic constructions -- specifically, according to Karmiloff-Smith, in order to avoid the ambiguity normally inherent in "même".

Vurpillot et al. (1971) and Vurpillot (1976) also noted children's adoption of the "same kind of thing" interpretation in the pair-comparison situation, while Donaldson and Wales (1970) and Webb et al. (1974) suggested that something along these lines also occurred with the term "different". That is, children who responded to a request for a different one from the target by selecting an identical one seemed to be taking "different" to mean "another one of the same class".

Two further implications follow from the proposal that the children in such studies were taking "same" to mean "same kind". Firstly, the important research question becomes not so much how well young children understand the word "same" itself, but why they interpret the test question in these task contexts to refer to qualitative sameness when for adults it is unambiguously perfect identity across all features that is

---

* This finding emphasises that care should be taken in interpreting the results of pair-comparison tasks in which the test question is something like "Are they the same?". Different forms of the question may elicit different answers. Cf. the footnote on pp. 94 and 96.

meant. Secondly, children do not make confusion errors simply because they take "same" to mean "alike in some respect(s)" instead of "identical in all respects". It is not <u>any</u> kind of sameness that will do! The precise criterion or criteria that a child will use in a given situation as the basis for deciding that stimuli belong to the same class will be defined for him by the stimulus context.

The influence of context has been commented on in the past (e.g., by Rosenblith, 1965; Garner, 1966; Garner & Felfoldy, 1970; Taylor & Wales, 1970; Gregson, 1975; Tversky, 1977). The present results reaffirm its importance, in this case for the particular "meaning" it can impart to individual stimuli or individual stimulus features. To illustrate: The subjects given the second stimulus set, in Experiments 4 to 6, were consistently highly successful in discriminating the mirror-image curved lines that they interpreted as "happy" and "sad" mouths. Yet they often failed to discriminate those very forms when they appeared as nose or eyes in a face. Again, subjects were able to differentiate the two sets of mirror-image diagonals that formed the Eyes of the third stimulus set. But when presented in isolation, such orientation and reversal transformations are notoriously confusable for young children. Clearly, the face context facilitated differentiation of those lines because of the meaning that they took on: they were not the same because one signified "happy" and the other "sad" or "cross". Deprived of that meaning, the isolated lines could be regarded as the <u>same kind of thing</u> -- both curved lines.

The combined notions that young children take the qualitative "same kind of thing" interpretation of the Experimenter's instructions, and that they attend to the "meaning" ascribable to stimuli or their features on the basis of their context, could thus perhaps help to explain some of the results reported by other investigators. As noted, confusion errors are commonly found for stimuli that differ by orientation, reversal, perspective, etc. (see Part 1, pp. 8ff). These sorts of transformations -- what Zusne (1970) calls "transpositional" operations -- do not alter the "essence" of the figure itself; it remains the same kind of thing. Moreover, young children are accustomed to ignoring such transpositional variation in the real world. (Two physically identical toys remain, for most purposes, identical if one of them is turned upside-down.) At the other extreme, differences of form or contour itself are rarely confused by young children (see pp. 13ff). These sorts of transformations -- called "topological" by Gibson and "transitive" by Zusne -- do change the essence of the figure, in a way that may mean it is no longer the same kind of thing.

Clearly at the moment these are speculative proposals that remain to be verified. But the tests should include asking the children not just whether items are the same, but why they are the same and what differences there are between unidentical stimuli that they decide are the same. Such questions have not often been systematically asked of children in past research on confusion errors; in the present case, their answers have proved highly informative.

A few final, broad remarks. It is obvious that the present study has very little to say about the <u>developmental</u> side of children's sameness judgments. Where age groups were examined separately, older subjects were generally found to make fewer confusions than younger ones. In children over about five-and-a-half years old, correct judgments predominated, or where errors did occur the child usually at least showed signs of attending to number of points of sameness along with kind of sameness. But no light is shed on what brings about this transition from the childlike mode of operating in terms of qualitative classification into the adultlike mode of deciding on the basis of quantitative physical identity -- whether it be changes taking place within the child's cognitive system and/or imposed from outside (for example, via education from adults), or whatever. The general age trend was complicated, moreover, by the anomalous results on Nose judgments in Experiments 4 and 5, where older children actually attained fewer correct discriminations than did younger ones. At the other end of the age-scale, some of the three-year-olds managed to achieve entirely correct performance; and those who made errors still came to the tasks with an already well-developed concept of sameness and difference. How they might have attained that level of competence in the first place is not elucidated.

The fact that such issues were not addressed in the present study is not to say that differences between age groups or changes over time in performance were regarded as unimportant to the questions of how children understand the term same, how they scan and process visual stimuli, and how they make decisions about sameness and difference. But the overall proportions of errors in general and their elimination with age were of only secondary interest here, because the present approach to these questions was through the types of confusion errors that occurred. The focus was therefore not on the ultimate achievement of error-free performance, but on the distribution of occurring errors among the stimulus attributes. The observed error-distributions generally were not found to differ between younger and older subjects in any major respect. Nevertheless, there were some suggestions of minor age-related variations. To determine conclusively why young children make confusion errors and what happens later to eliminate them, obviously the details of developmental changes must be investigated more rigorously than was attempted here.

The degree of competence that even the youngest subjects brought to the tasks in general, and the high success rate throughout the final experiment of the series in particular, serve to remind us of the mundane but nonetheless important point that laboratory-style studies can mislead; they may sometimes mask the true extent of the young child's capacities and provoke errors by artefact (see, e.g., Bryant, 1974; Donaldson, 1978). As Gelman and Gallistell (1978) emphasise, it is as important to consider what the child can do as what he fails in. The experimental psychologist's focus on confusion errors should not blind one to the fact that young children can deal accurately with sameness and difference -- and indeed must do so constantly as they get about in the world.

A comment by Schaller and Harris (1974) serves aptly to conclude this discussion. They suggest that the term "error" may be an inappropriate one for children's judgments of unidentical stimuli as the same. From the present results, too, it is clear that the subjects were not failing -- in the sense, for example, of failing to detect a difference. Rather, as Schaller and Harris point out, the children were taking active decisions about "when a difference makes a difference". While the outcomes of those decisions may not be what an adult would expect, they do not properly constitute errors; within the child's own system, they are legitimate. They reflect not so much misunderstandings of the situation as unconventional understandings of it. In 1972, Goodnow wrote that it is not enough for the developing child, in his intellectual pursuits, simply to enlarge his repertory of linguistic, perceptual and cognitive capacities. He must also acquire the "tricks of the trade", that is, he must learn if and when to apply a particular skill and be able to evaluate his success or failure. The present study reinforces Goodnow's point, for by age three or four years the child already has considerable perceptual and linguistic mastery of sameness and difference. At that level of competence, his job then becomes largely to learn under which circumstances a particular decision-strategy is appropriate or inappropriate -- to learn to conform to adult convention.

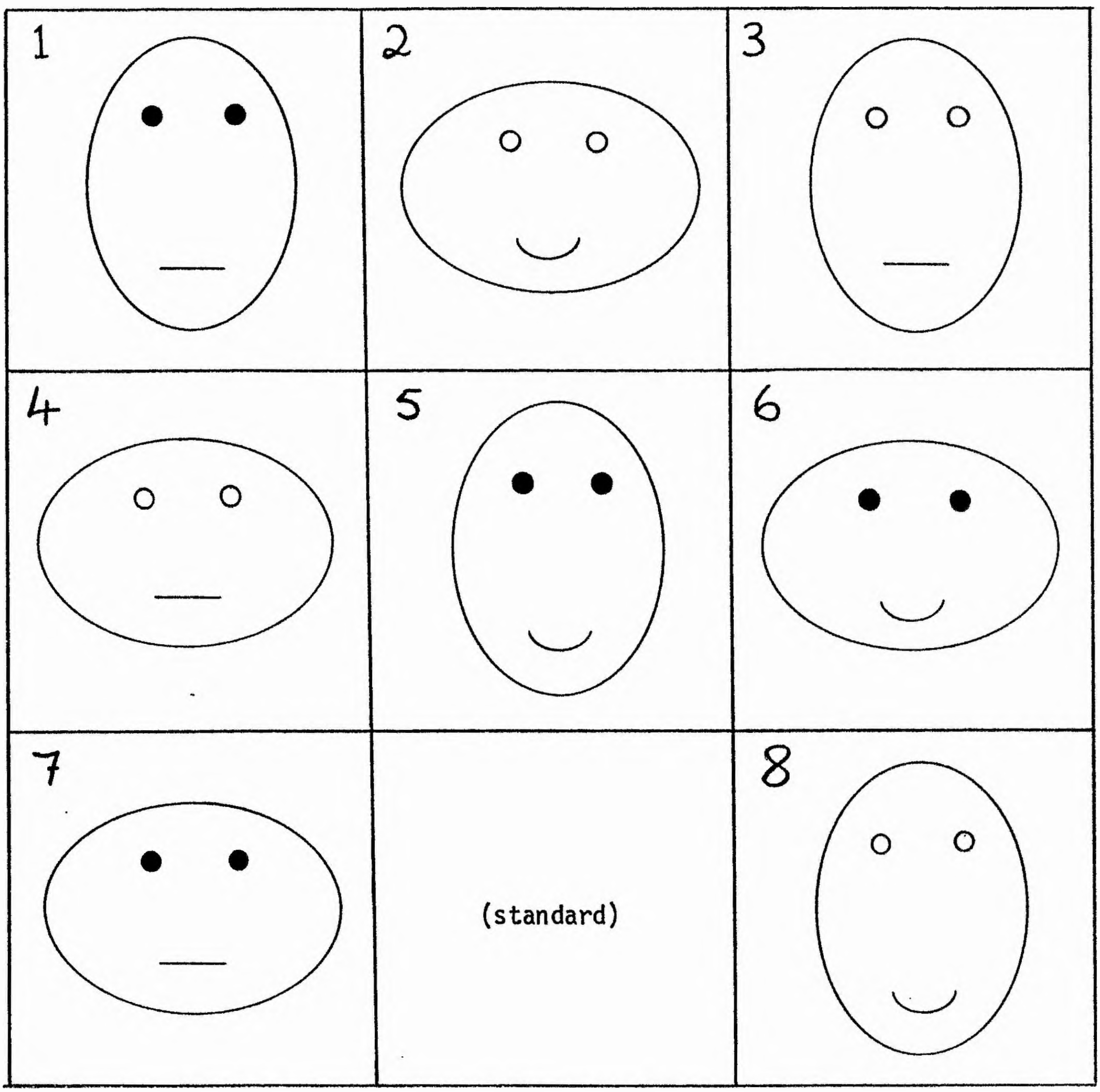
APPENDIX 1

Figures illustrating the three stimulus sets
employed in the present study

Appendix 1

Figure 1

First stimulus set -- used in Experiments 1, 2, 3 and 7.

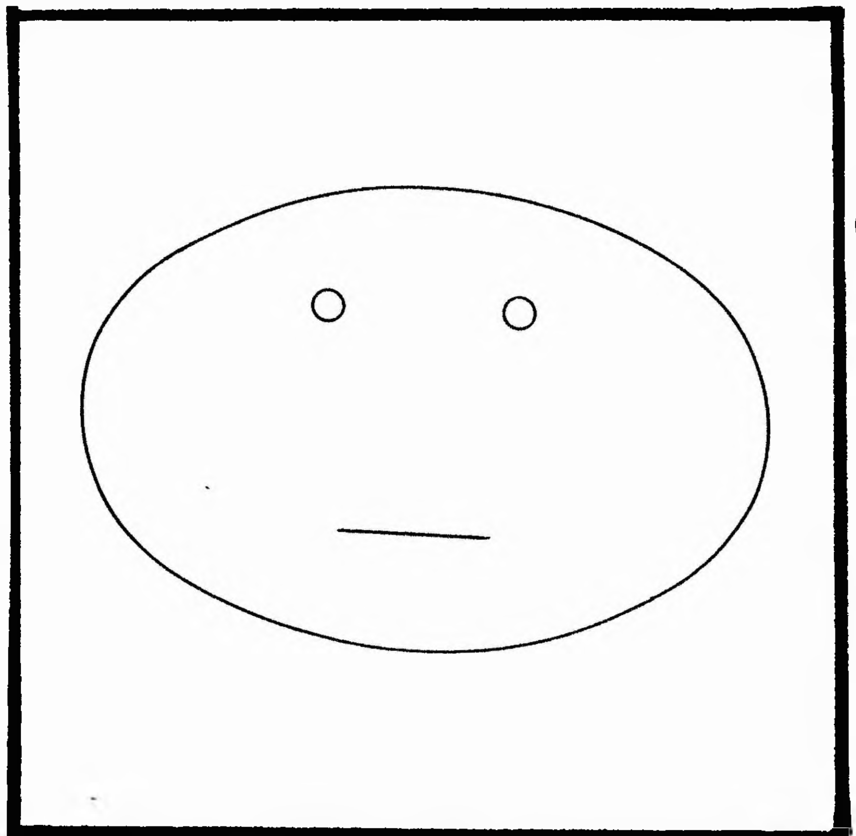


Stimulus composition followed that of Taylor (1973) and Wales (pers. comm., 1974). The above array illustrates the Framed set (cf. Taylor).

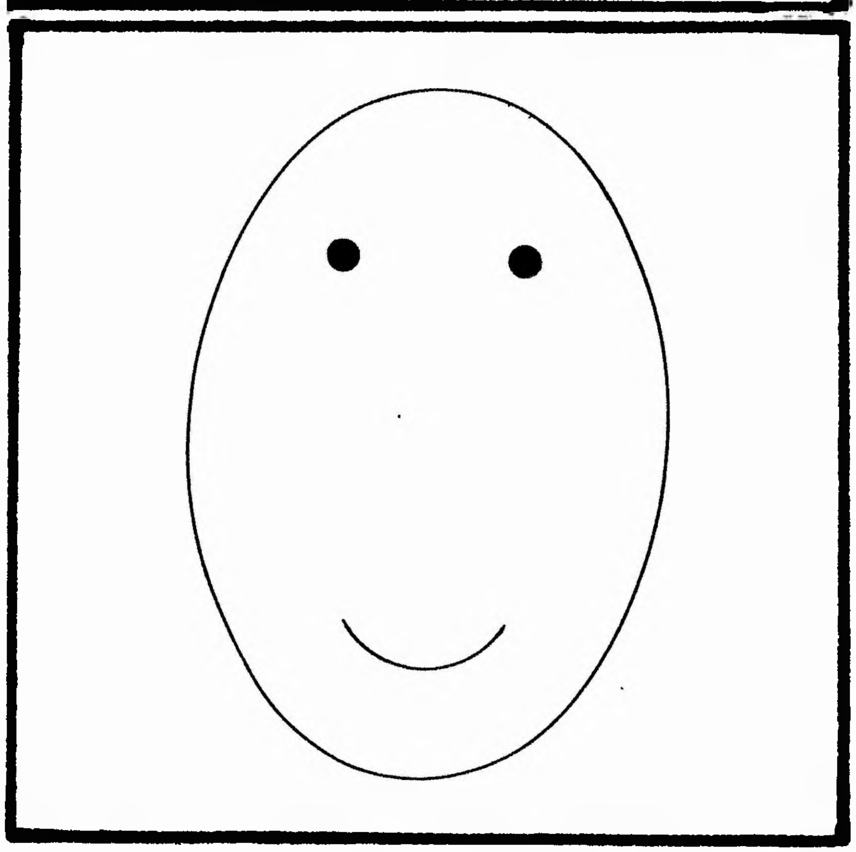
Appendix 1

Figure 2

First stimulus set: Examples of two of the faces actual size. (Between them, they illustrate both levels of all three attributes used in the construction of the eight faces in the set.)



SET 1
(FRAMED)
FACE
4

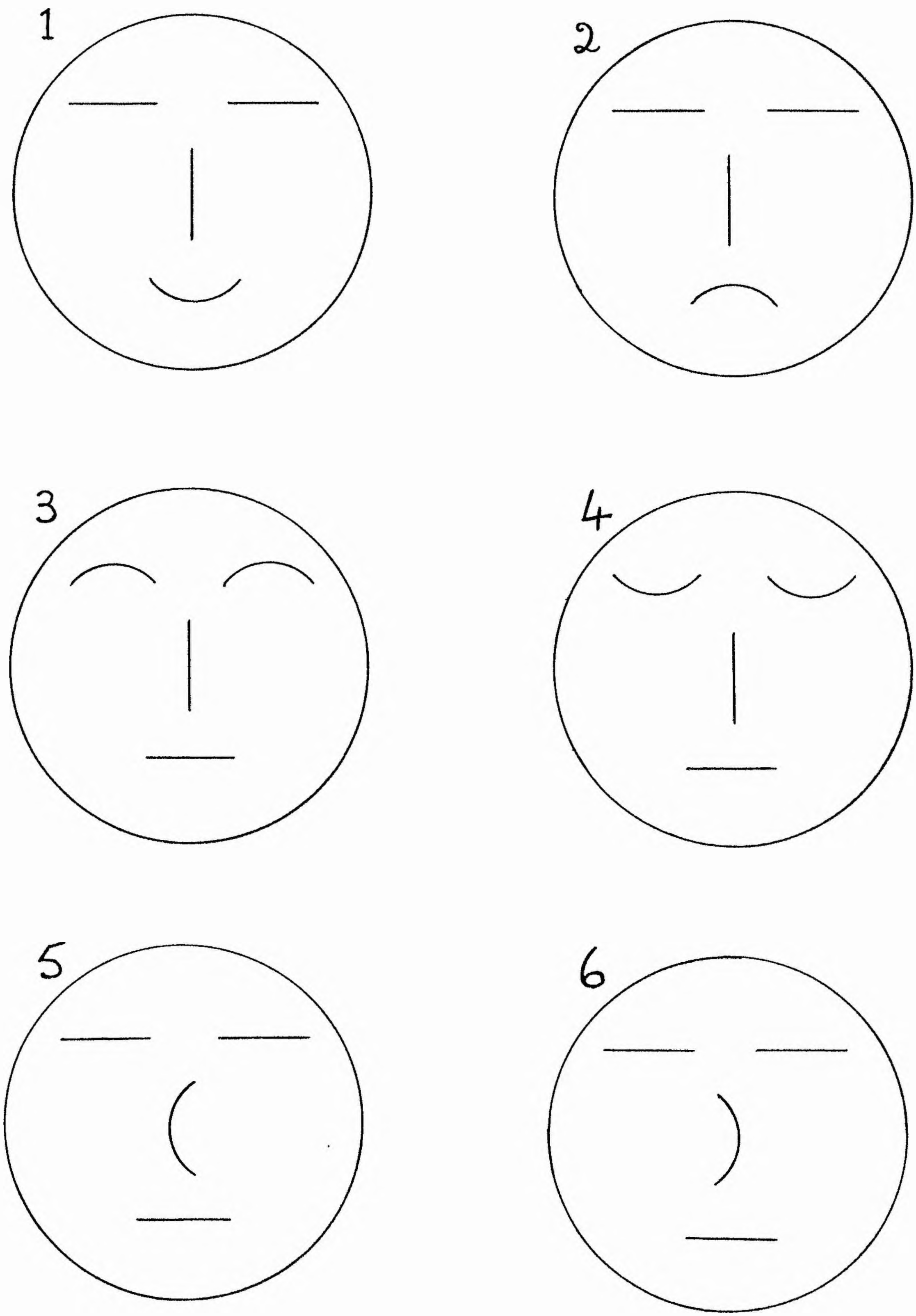


SET 1
(FRAMED)
FACE
5

Appendix 1

Figure 3

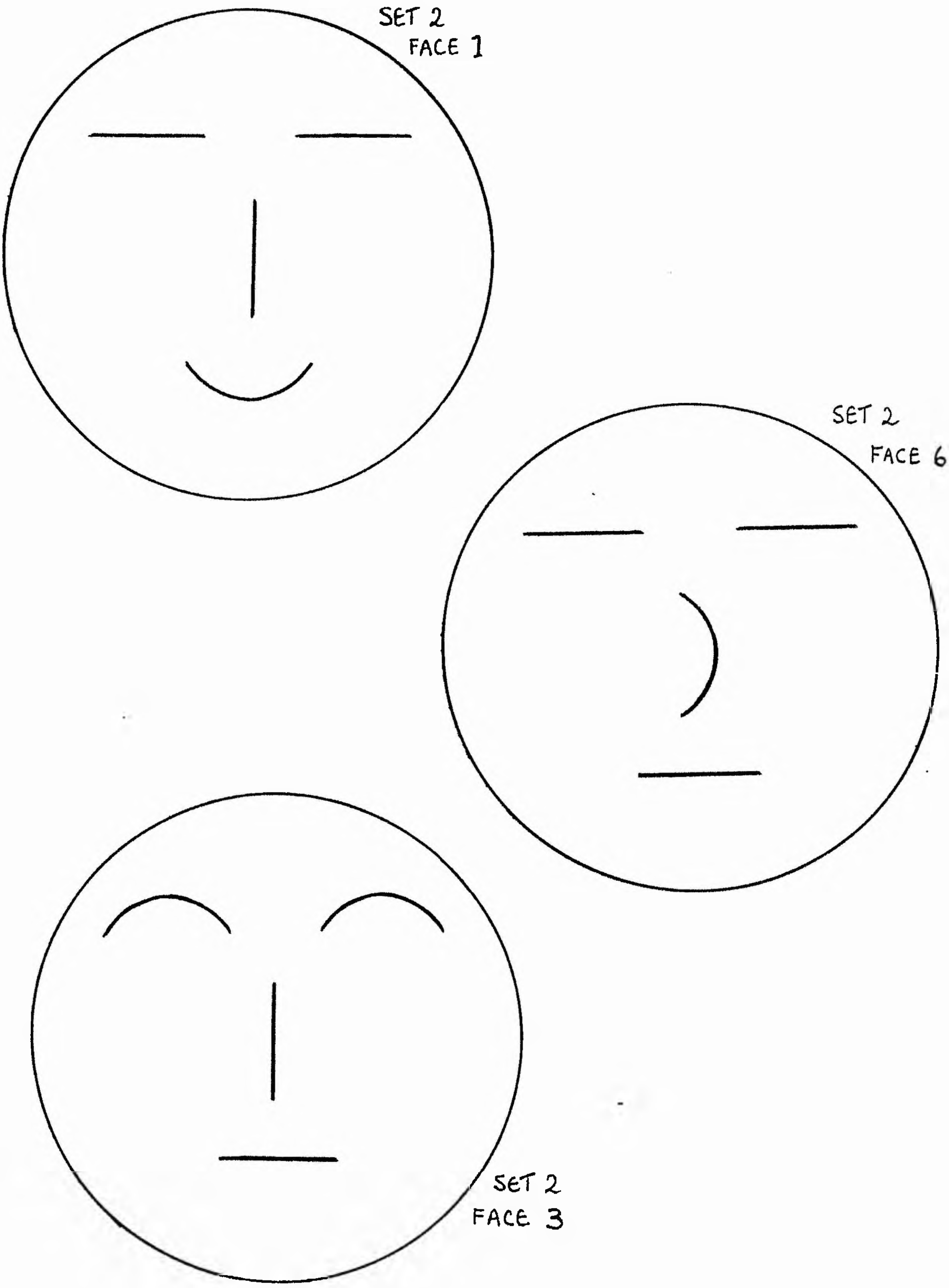
Second stimulus set -- used in Experiments 4, 5, 6 and 7.
(The same three variants make up each of the attributes)



Appendix 1

Figure 4

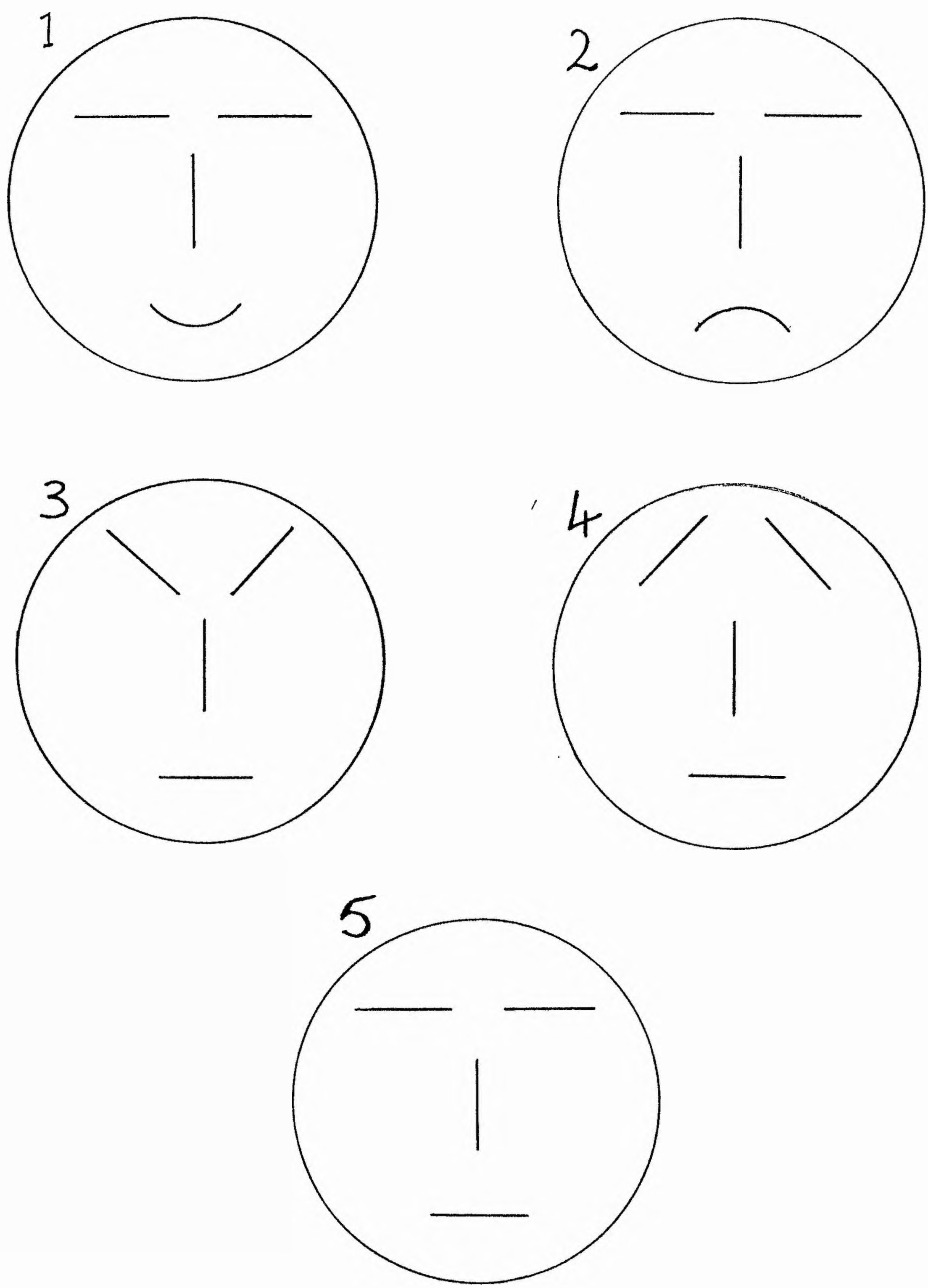
Second stimulus set: Examples of three of the faces actual size.



Appendix 1

Figure 5

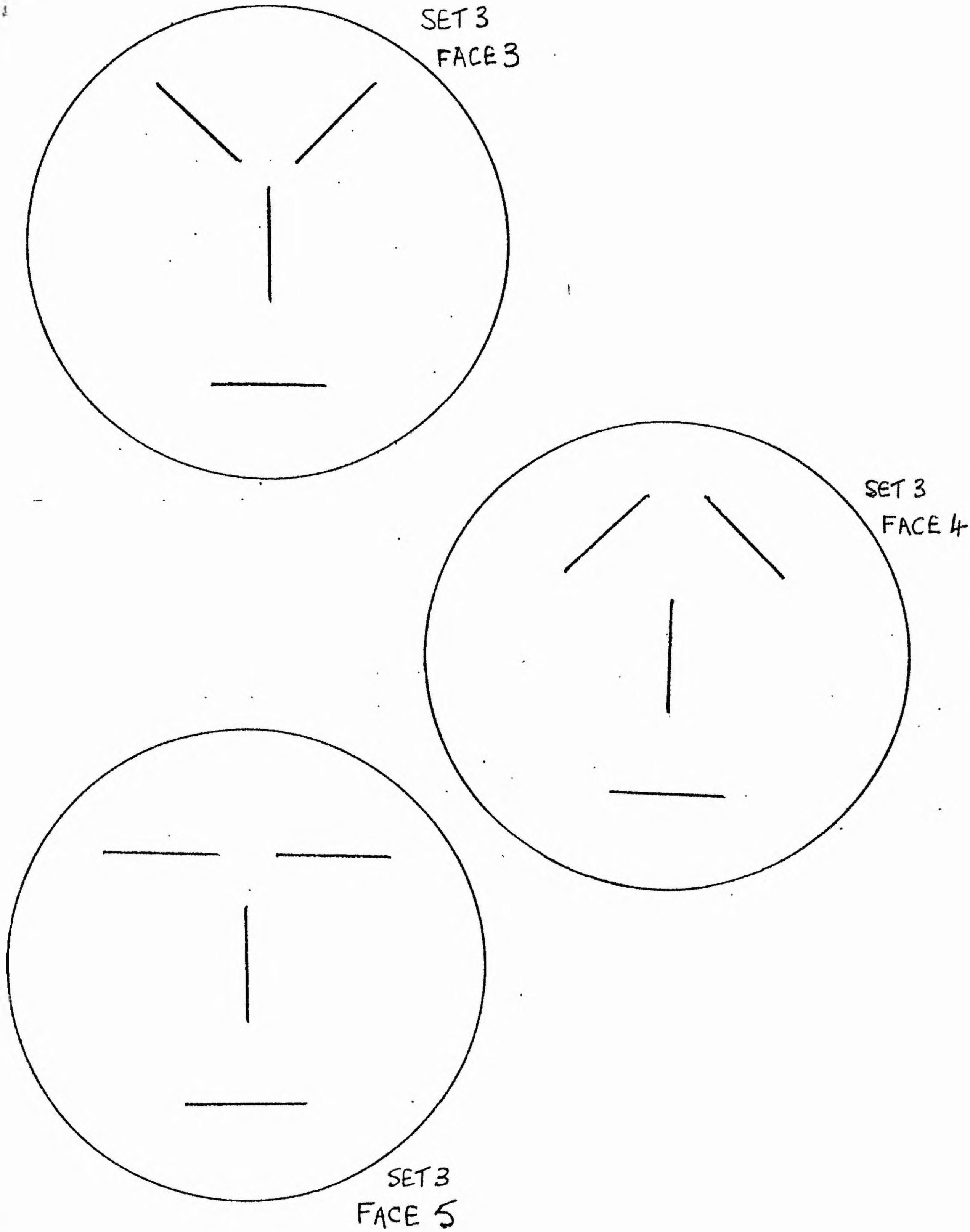
Third stimulus set -- used in Experiment 8.



Appendix 1

Figure 6

Third stimulus set: Examples of three of the faces actual size.



APPENDIX 2

SUPPLEMENTS TO EXPERIMENT 1

Section I

Comparison between Experiment 1 of the present study and Experiment III of Taylor (1973): Individual subjects' patterns of confusion errors

In Table XVI of his dissertation (1973), Taylor gives the numbers of subjects who followed particular patterns of confusions among the three attributes Shape, Eyes and Mouth -- i.e., the numbers who chose matches on the basis of a consistent "order of (perceived) similarity" among the attributes. He lists six such possible orders. For one-attribute confusions, these are: M > E > S, M > S > E, E > M > S, E > S > M, S > M > E, S > E > M (where ">" = "is/are more similar than" = "is/are more often confused than"). The corresponding orders for two-attribute confusions are: ME > MS > ES, MS > ME > ES, ME > ES > MS, ES > ME > MS, MS > ES > ME, ES > MS > ME. For the present study, however, it was possible to say only that particular individuals made the most confusions, e.g., with Mouth (and fewer with Eyes and Shape <u>equally</u>), or most with Mouth and Eyes <u>equally</u> (and fewer with Shape). Finer distinctions of order among the attributes were usually not discernible in the present Experiment 1.

For comparison with Taylor's pattern of results, these data are presented in Table A2-1 (next page). Taylor's six possible orders have here been collapsed into three, to show simply the numbers of subjects who confused Mouth most, or Eyes most, or Shape most. Three new categories have been added to take account of those subjects in the present study who confused <u>two</u> attributes equally often. Also shown are the

TABLE A2-1 (Experiment 1)

Order of perceived similiarity among attributes: Number of subjects most confusing each attribute or combination of attributes.

Attribute(s) most often confused	Taylor's Sample ① (Framed Set)	Present Sample ② Framed Set	Present Sample ② Unframed Set
Mouth	32	3	3
Eyes	9	6	6
Shape	7	1	0
Mouth & Eyes equally	N.A.	0	2
Mouth & Eyes equally	N.A.	0	0
Eyes & Shape equally	N.A.	2	2
unclassifiable	9	6	5

Key:

① Compiled on the basis of Table XVI in Taylor's dissertation (1973).

② Data for first responses only.

N.A.= not applicable. Three categories here are not applicable to Taylor's data from his Table XVI. All of his subjects fit into the first three categories, except for the nine omitted from his Table XVI, who are "unclassifiable".

numbers of "unclassifiable" subjects: those whose confusion errors were too infrequent or too inconsistently distributed over the three attributes to allow categorisation.*

Shape was the attribute confused by the fewest subjects in all cases. But whereas most of Taylor's subjects confused Mouth above all, the tendency in the present sample, with either stimulus set, was to confuse Eyes most of all. This merely confirms what we have already seen from the analyses of pooled response patterns for Experiment 1 presented in the main text.

---

* In this experiment, Taylor ran a total of 57 subjects. Forty-eight are accounted for in his Table XVI, so presumably nine were omitted from this analysis -- a rate of 16% unclassifiable. In his Table XVII, however, he states that only 33 children showed a consistent order, while seven were inconsistent, and 17 were "not relevant" (because they made too many perfect matches or two-attribute matches). This would appear to give in reality a rate of 42% unclassifiable. Cf. the rates of 28% and 33% (five and six out of 18) unclassifiable in the present sample, shown in Table A2-1.

Section II

Differential choice frequencies among the eight stimulus faces of Experiment 1: Comment on response biasses, and supplementary test

Why should certain of the faces be chosen more often than others as match to _any_ standard? Recall that the order of the stimulus array was constant over all subjects and all conditions. Figure 1 of Appendix 1 displays the eight stimulus faces in this fixed array. Faces numbered 1, 2 and 3, which were among the least often chosen as match to any standard, formed the top row of the array. Could it then be that the children simply chose those faces nearest them in the array and overlooked the top row? (A bias towards the right-hand side of the array would have to be added too, to explain the relatively few choices of face 4, at the left of the middle row.)

Such a position preference seems unlikely. For one reason, Taylor's subjects apparently did not exhibit a position preference in his corresponding matching-from-sample task with these schematic faces. Although he does not state so directly, there are two grounds for supposing this: (1) Where position biasses arose elsewhere in his series of experiments, he notes them explicitly; but he makes no direct mention of them at all in this particular context. (2) He does note certain asymmetries in the frequencies of choices, referring to cases where stimulus A is chosen more often as match to stimulus B than B is chosen as match to A (Table XXII of his dissertation, 1973). But these asymmetries exhibited no consistent pattern, and in any case do not correspond with the preferences found in the present Experiment 1. So there is no _a priori_ reason to expect a position bias.

Secondly, several adults were informally confronted with the fixed stimulus array and were asked, "Which face do you like best/least?". Faces 5 and 8 were preferred, and face 4 was least liked -- a pattern that roughly paralleled the children's choices in the matching-from-sample task.

To check more formally on the possible influence on the children's choices of global preferences among the faces, as opposed to a real position bias, a supplementary test was conducted.

The subjects were 22 children from the first-year class of a primary school in St. Andrews and from an adjoining preschool nursery class. They were divided into four groups, according to whether they received Framed or Unframed stimuli first and whether they received the original stimulus array or a new, "random" one:

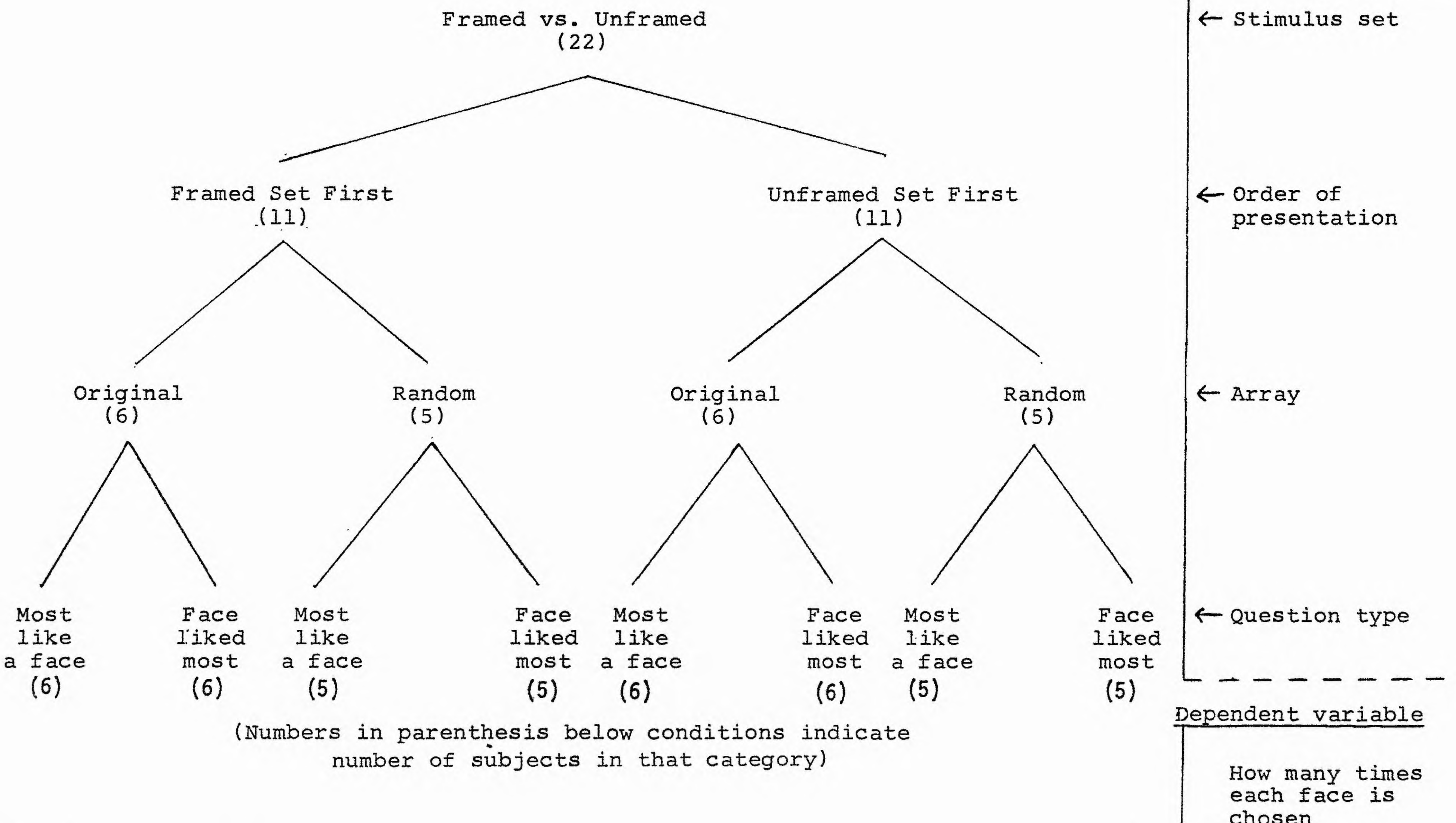
Group I-O: n = 6; ages 4;3 - 5;7, mean 4;9 years
Group I-R: n = 5; ages 4;2 - 5;7, mean 4;10
Group II-O: n = 6; ages 4;3 - 5;4, mean 4;10
Group II-R: n = 5; ages 4;4 - 5;5, mean 4;10.

The experimental design is summarised in Figure A2-1 (next page). The stimuli were those used in Experiment 1: the eight faces of the Framed set and eight cut-out faces of the Unframed set. Children in Groups I received the Framed set first, and the Unframed about a week later; the order of presentation was reversed for Groups II. So as in Experiment 1 proper, each child acted as his own control with respect to the two stimulus sets.

A new stimulus variable was added: Original versus Random array. For Groups I-O and II-O, the order of the faces in the array was the same

Figure A2-1

Summary of design of test for face preferences



(Numbers in parenthesis below conditions indicate number of subjects in that category) as in Experiment 1 (see Figure 1 of Appendix 1), constant over all subjects in these groups. Groups I-R and II-R received a new, "random" array. In the original array, recall, faces 1, 2, 3 and 4 (the least often chosen as matches to any standard) appeared on the top and middle rows, while faces 5, 6, 7 and 8 were on the middle and bottom rows. To control for position biasses, in the new array faces 1, 2, 3 and 4 now appeared in the middle and bottom rows, and faces 5, 6, 7 and 8 in the top half. Within the limits of this proviso, the locations of the faces in the new array were determined randomly, the order differing across children in Groups I-R and II-R and across test sessions with the same child.

Subjects were tested individually. In each session, the appropriate stimulus set and array was laid before the child, who was asked, (1) "Which one looks most like a face?" and (2) "Which face do you like the most?". The order of the two questions was randomised across subjects and sessions.

The general pattern of results is shown in Table A2-2. In overall terms, faces 5, 6 and 8 were chosen the most often. This accords with the distribution of choices as match to any standard that emerged in Experiment 1 proper (cf. Table 12, p. 58), except that the high choice frequency observed earlier for face 7 has now been superseded by that of face 2. Little information was gained by further breakdown of the choice frequencies into sub-conditions (e.g., comparing results on Original against Random array within the sub-conditions Framed set plus Question "Which is most like a face?", etc.). Fisher tests on the numbers of choices accruing to faces 1 to 4 versus 5 to 8 in such subgroups gave no significant values in any instance. There were no apparent differences

TABLE A2-2.

Distribution of choices (%) among the eight faces. The highest scores (≥ 15%) are ringed, to show more clearly which of the eight faces were the most often chosen in each sub-condition.

Experimental Condition*		Face Number 1	2	3	4	5	6	7	8
Array									
Original	(n = 12)	4	8	2	4	33	29	6	13
Random	(n = 10)	3	13	4	0	30	15	10	25
Question-type									
Most like a face	(n = 22)	5	5	2	2	41	18	2	25
Face liked most	(n = 22)	2	16	5	2	23	27	14	11
Stimulus set									
Framed	(n = 22)	5	11	5	2	25	23	9	20
Unframed	(n = 22)	2	9	2	2	39	23	7	16
Overall	(n = 22)	3	10	3	2	32	23	8	18

* Within each condition presented here, all other conditions are confounded; e.g., each of the array levels Original and Random cuts across both levels of stimulus set, question-type, and order of presentation.

between the two levels of any condition, nor any interactions among conditions.

The absence of a difference between the Original and Random array conditions argues against the position-preference view suggested earlier as a possible explanation of the uneven distribution of choices over the eight faces. Faces 5, 6 and 8 between them account for 70% of the choices in the "random" array -- almost double the rate expected on an even-distribution basis -- despite the fact that these stimuli were now located near the top of the display, farthest from the child. Face 2 ranked next. Notice that all four of these "most popular" faces share one feature in common: the Mouth variant is always the curved form resembling a "smile". It may therefore be that the smile was an important variable in attracting children's choices.* In fact, two of the subjects (one each from Groups II-O and II-R) who chose face 7, with the straight Mouth, as the face most liked volunteered reasons: "It's the funniest-looking face" and "It doesn't go with the other" -- where "the other" referred to the child's most-like-a-face choice (face 8 for both these subjects).

In any case, the absence of position bias, together with the fact that the most popular faces here fit reasonably well with the faces most often chosen in Experiment 1, suggest that the children's responses in the matching-from-sample task may have been partly governed by subjective preferences among the faces. Consider, for example, a child who chooses

---

* At the time, this was a tentative suggestion. The results of the next experiments, however, will be seen to point in this direction, too. The precise significance of Mouth and the smile will be directly examined in their own rights in later parts of the study.

matches on the basis of shared Shape and Eyes, and confuses Mouth. Given face 1 as the standard, he has two possible Shape-plus-Eyes-sharing alternatives: faces 1 and 5. It is possible that he regards these as equivalent in terms of their goodness of fit to the standard; but he may then decide on face 5 rather than face 1, simply because he likes it better. The possible role of such subjective preferences in multiple-choice tasks of this nature deserves further study.

APPENDIX 3

EXAMPLES OF SUBJECTS' JUSTIFICATIONS FOR THEIR SAMENESS JUDGMENTS

<u>Contents</u> <u>page(s)</u>

KEY TO ABBREVIATIONS (ETC.) USED IN JUSTIFICATION-TRANSCRIPTS

M, S, E, N	Mouth, Shape, Eyes, Nose
C	Child (subject)
( )	Anything enclosed in brackets is Experimenter's addition, not uttered by subject
(p.)	Subject pointed to stimulus or part of it
e.g., (p. 4)	-- Subject pointed to face 4
(p. M on 56)	-- Subject pointed to Mouths on faces 5 and 6
(Q)	Experimenter asked a further question. E.g., if child said, "They're <u>all</u> the same" in triads task, Experimenter asked for "the <u>two most</u> the same", etc. Or if child gave a wrong "same" judgment but then explained how the pair <u>differed</u>, Experimenter asked more about sameness and <u>questioned</u> that if there was something different, could they be "just the same". Etc.
---	Child offered no spontaneous justification and Experimenter did not press for one
...	Child broke off speech (usually to point at something instead of continuing to comment verbally)
(?)	The word immediately following the bracketed question-mark is not certain, since the child spoke it unclearly and did not clarify it
<u>underlining</u>	Word(s) underlined in child's utterance were spoken with heavy stress
=	Against <u>triad</u> trial, denotes a triad in which all three faces <u>shared</u> the same level of the subject's dominant matching attribute, and which elicited an "<u>all</u> the same" response (see main text, p. 174)
*	Against <u>pair-comparison</u> trial, denotes a pair that received a wrong pair-comparison judgment -- a false "same" response for an unidentical pair, or a false "not the same" response for an identical pair. (All trials not marked by the asterisk were correctly judged.)

The numbers given in the columns headed <u>TRIAD</u> or <u>(CHOSEN) PAIR</u> correspond to the numbers assigned to the various stimuli within each set. These numbers are shown against their faces in Appendix 1, Figures 1, 3 and 5. (For ease of reference, rough drawings of the stimuli with their numbers have been provided at the beginnings of the appropriate lists of examples.)

EXCERPTS FROM RESPONSE SHEETS: TRIADS (Experiment 2)

(a) Shape-matchers

	TRIAD	CHOSEN PAIR	SHARING	JUSTIFICATION	E's NOTES
Boy, age 4;9 Framed set	587	58	SM	Two smiling faces	
	345	35	S	They're both the same	cross/grumpy = straight Mouth
	167	67	SE	Except that one's cross (7) and that one's smiling	
	126	26	SM	They've both got the same mouth	
	457	47	SM	They're both grumpy faces	
	368	38	SE	Except... (p.M). They're the same shape	
	268	26	SM	But that one's got black eyes (p. 6 then p. 2)	
	158 =	15	SE	They've both got black eyes. They're all the same.	= 158 all share Shape
	236	26	SM	---	

	TRIAD	CHOSEN PAIR	SHARING	JUSTIFICATION	E's NOTES
Boy, age 3;4 Unframed set	578	58	SM	I don't know. (Q; C just nodded)	
	345	35	S	I don't know. (Q)	
	167	67	SE	That's a fat face and that's a fat one	
	126	26	SM	Got the same mouth and eyes ←	Eyes wrong
	457	47	SM	They two have the same mouth	
	368	38	SE	---	
	268	26	SM	(p. M)	
	158 =	58	SM	(p. all 3 faces in turn) (Q) (chose 58, refused comment)	= 158 all share Shape
	236	26	SM	They two's got round mouths	

EXCERPTS FROM RESPONSE SHEETS: TRIADS (continued)

(b) Mouth-matchers

	TRIAD	CHOSEN PAIR	SHAR-ING	JUSTIFICATIONS	E's NOTES
Girl, age 6;4 Framed set	128	28	EM	They're both happy	
	237	37	M	They both have straight mouths	
	148	14	M	They both have straight mouths	
	258 =	58	SM	They both have the same shape and they're both happy. They're all happy but these two are best the same.	= 258 all share Mouth
	367	37	M	---	
	134 =	13	SM	These two both have the same shape, but they're all the same.	= 134 all share M
	568 =	58	SM	They're all the same, they're all happy. But these two are best the same.	= 568 all share M
	135	13	SM	They're the same shape	
	257	25	M	They're both happy	
	348	34	EM	Straight mouths, and they both have the same eyes	

	TRIAD	CHOSEN PAIR	SHAR-ING	JUSTIFICATIONS	E's NOTES
Boy, age 5;3 Unframed set	128	28	EM	They've got the same eyes and the same mouth	
	237	37	M	They've got the same mouth	
	148	14	M	Same mouth	
	258	58	SM	The same mouth	
	367	37	M	They've got the same mouth	
	134	34	EM	The same mouth	
	568 =	56	EM	All of them. (Q) (Chose 56:) They've got the same eyes	= 568 all share M
	135	13	SM	But...(p. 1 & 5, Eyes) same eyes	
	257	25	M	They've got the same mouth	
	348	34	EM	They've got the same eyes and the same mouth. They've all got the same eyes but not the same mouths.	

EXCERPTS FROM RESPONSE SHEETS: TRIADS (continued)

(c) Eyes-matcher

	TRIAD		CHOSEN PAIR	SHARING	JUSTIFICATION	E's NOTES
Girl, age 3;9 Unframed set	348	=	34	EM	(p. all 3 faces in turn) (Q) (Chose 34:) They're two sads.	= 348 all share Eyes
	567		56	EM	They're funny (= M on 56) and that's sad (= M on 7).	funny = happy i.e., curve M sad = straight Mouth
	147		17	EM	---	
	256		56	EM	Two of them are funny and one's sad.	
	356		56	EM	Funny (p. M on 56). Sad (p. M on 3).	
	136		16	E	---	
	248	=	28	EM	(p. all 3 faces) (Q) (Chose 28) Two funny and one sad (M on 28 vs 4).	= 248 all share E
	145		15	SE	---	
	238	=	28	EM	(p. all 3 faces) (Q) (Chose 28) They're funny (M on 28) and that's sad (M on 3).	= 238 all share E

(d) Mixed-matcher

	TRIAD	CHOSEN PAIR	SHARING	JUSTIFICATION
Girl, age 5;11 Framed set	456	56	EM	They're both smiling and two black eyes.
	378	37	M	They've got lines across the mouth.
	125	15	SE	They've both got black eyes.
	168	68	M	They've both got a smile on the face.
	235	25	M	They've both got a smile on the face.
	578	58	SM	They've both got a smile.
	345	34	EM	They've both got white eyes and a line across the mouth.
	167	17	EM	They've both got black eyes and a line across the mouth.
	126	16	E	They've both got black eyes.

EXCERPTS FROM RESPONSE SHEETS: PAIR-COMPARISON (Experiment 3)

	PAIR	SHAR-ING	JUSTIFICATION	E's NOTES
Shape-matcher girl, age 4;3 Framed set	46	S	The eyes are not the same but the face is and the mouth isn't.	Not clear if face = just Shape, or whole stimulus. In coding "Type of Reference", classed as whole stimulus.
	47	SM *	The faces are but the eyes aren't (the same)	
	48	E	---	
	56	EM	The eyes are the same.	
	57	E	Not that one (p. M), not the mouth.	
	58	SM *	Not the eyes. (Q) The mouths are right. (Q) They're the same.	
	67	SE *	---	
	68	M	The eyes are not the same.	
Mouth-matcher boy, age 5;2 Unframed set	13	SM *	The two eyes and the flat mouths are the same. One's got eyes not coloured and that one is coloured.	Note this C perfectly happy to call a pair "just the same" while explicitly admitting some difference. Insisted on "same" even when asked how they can be the same when they are different.
	14	M *	They're the same but different. (Q) They're the same, both are egg-shaped. But one is a standing-up egg and the other is a fallen-down egg. (Q) They're the same.	
	15	SE	No, but something's the same. They're both standing-up eggs and both got coloured-in eyes.	
	16	E	---	
	17	EM *	That one's a standing-up egg and that one's a falling-down egg. They both got flat mouths and they both got coloured-in eyes.	
	18	S	One's sad and one's happy.	
	23	E	Not the same. But they've both got no coloured-in eyes.	
	24	SE	But they're both fallen-down eggs.	
	25	M *	They're both happy. But one hasn't got coloured-in eyes and they've got different shapes.	
	26	SM *	(C kept changing his mind: first not same, then same, not same, finally same:) But this one's got no coloured-in eyes and this one has (coloured eyes).	

EXCERPTS FROM RESPONSE SHEETS: PAIR-COMPARISON (continued)

	PAIR	SHAR-ING		JUSTIFICATION	E's NOTES
Eyes-matcher	17	EM	*	(p. M) Both got black eyes	
boy, age 5;1	18	S		One's got a flat face and one's happy and they've both got different eyes	flat face = straight Mouth
Unframed set	23	E	*	(long time before deciding; refused reason)	
	24	SE		---	
	25	M		But they've both got smiling faces	
	26	SM		One's got black in the eyes and one's got no black	
	27	S		No, but they've both got the same shape. But they've got different faces, different eyes.	
	28	EM	*	Yes, but they're both different shapes	
	34	EM	*	---	

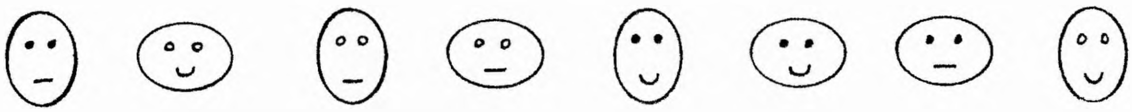
	PAIR	SHAR-ING		JUSTIFICATION	E's NOTES
S+M-matcher	37	M		That's a long mouth and that's a long mouth, but the shape is wrong	
boy, age 4;0	38	SE		A long mouth and a round mouth	
	45	nil		It's a round mouth and a straight mouth	
Framed set	46	S		---	
	47	SM	*	That's a long one and that's a long one (p. M)	
	48	E		A long mouth and a square mouth	
	56	EM		It's only the same mouth. It isn't the same head	← C makes explicit his criteria!
	57	E		A long mouth and a square mouth	
	58	SM	*	Square mouth and square mouth	square = curved Mouth
	67	SE		A long one and a square one (Re. M)	

EXCERPTS FROM RESPONSE SHEETS: PAIR-COMPARISON (continued)

	PAIR	SHAR-ING	JUSTIFICATION	E's NOTES
Mixed-matcher girl, age 5;4 Framed set	34	EM *	They're the same but one's fat and one's thin.	
	35	S *	One's sad with white eyes and one's happy with black eyes. They're both thin heads. (Q) They're both the same.	
	36	nil	One's got black eyes and a fat face and he's smiley and the other's got white eyes and a thin face and he's sad.	
	37	M	One's thin with a sad face and one's a fat face and it's not got black eyes.	
	38	SE *	One's sad and one's happy. (Q) They've both got thin faces, they're the same.	
	45	nil	---	
	46	S	One's got black eyes and a smiley face and one's got white eyes and a sad face.	
	47	SM *	---	
	48	E	One's thin and one's fat, but they've both got white eyes.	
	56	EM *	They're the same but one's fat and one's thin.	

	PAIR	SHAR-ING	JUSTIFICATION	E's NOTES
Correct-matcher girl, age 4;6 Framed set	15	SE	They're black eyes, that's a smile and that's sad.	
	16	E	It's black eyes and a smiley face, black eyes and a straight face.	straight face = straight Mouth
	17	EM	Black eyes, black eyes. Straight mouth, straight mouth. But they're not the same because the body is not the same.	body = Shape
	18	S	It's got no eyes (= 8). That's a sad mouth (= 1). But the body's the same.	
	23	E	A straight face and eyes and a round body (= 3) and a smiling face and round eyes (2).	
	24	SE	That's a straight sad mouth and no eyes (4). But it's the same body.	
	25	M	---	
	26	SM	It's got no bright eyes (= 2). That's a smile, black eyes (6).	
	27	S	The same head. There's a smiley mouth and white eyes (2), a straight face, black eyes (7).	

EXAMPLES OF INDIVIDUAL JUSTIFICATORY RESPONSES
(Experiments 2 & 3: Triads & pair-comparison)



1 2 3 4 5 6 7 8

EXAMPLE	TRIAD-CHOSEN PAIR	SHAR-ING	JUSTIFICATION	C's AGE
1.	136-13	SM	They've got the same everything except they eyes. They've got the same shapes.	4;4
2.	123-13		They've both got a round head, straight face, and the same eyes, nearly. (round head = vertical S; straight face = straight M)	5;9
3.	123-13		They've nearly the same face. The mouth is the same but the eyes are not.	6;4
4.	347-47		They're all alike but these two (47) are best. They've got the same shaped faces and the same eyes. (Eyes wrong!)	6;4
5.	147-47		I don't know, they're all some bit alike. (Refused to choose for a long time, then refused reason)	6;11
6.	267-26		They're all the same. (Q) All of them. (Q) They all go together. (Q) (C chose 26:) They've both got sort of round faces, both got round eyes and a round mouth.	3;5
7.	267-26		They're all the same. I'm going to take the two smiley faces.	4;9
8.	258-58		Maybe all of them because they're all smiling. But these two (58) are best because of their round heads (= vertical S).	5;5
9.	568-58		Because of their straight heads (= horizontal S) and they're both smiling. But they're really all the same.	"
10.	258-58		They've got the same shape and mouth. They've all got the same mouth. But that one (5) hasn't got the same eyes.	5;5
11.	135-15	SE	All same. (Q) (Chose 15:) They're a little bit not the same. That one's smiling and that one's not.	4;8
12.	135-15		They've both got black eyes and... just black eyes.	5;11

EXAMPLES OF INDIVIDUAL JUSTIFICATORY RESPONSES (cont.)

EXAMPLE	TRIAD-CHOSEN PAIR	SHARING	JUSTIFICATION	C's AGE
13.	567-67	SE	Because these two differ on just one thing (p. M). All the other two's differ on two things. (Not True: 56 share EM!)	6;4
14.	248-24		They're all the same. (Q) They all don't go. (Q) They all are (the same). (Q) These two (24) are the same in heads and eyes but not the mouths. These two (28) are smiling, they're the same except the heads. These two (48) are the same except the mouth and heads. (Q) (Finally chose 24 but refused reason.)	"
15.	248-28		They've all got the same eyes. But these ones (28) have got the same eyes and the same mouth.	5;2
16.	278-28	EM	They've got smiling faces and the eyes are the same, but not the head. This one (7) doesn't go because these two things don't go. (p. E & M).	6;2
17.	147-17		All of them. (Q) (Chose 17:) They've got the same eyes and mouths, except for the shapes.	4;4
18.	348-34		They've got the same mouth and eyes. But they all have the same eyes. But those two (38) don't have the same mouths.	5;5
19.	148-18	S	They've both got round heads (= vertical S). But if they all had straight mouths, they'd all be the same.	5;5
20.	237-23	E	They have the same face except for the mouth.	4;4
21.	167-16		They two got black eyes (16). (Then re. 67:) They don't fit the same because they've both got... (p. horizontal S).	5;4
22.	348-48		They've both just got white eyes.	5;11
23.	478-78	nil	They've nearly got the same kind of eyes.	5;3
"Unclassifiable" responses in "Type of Reference" coding:				
24.	168-18	S	The mouth's changing to the other mouth.	6;2
25.	458-48	E	The mouth is covered up.	4;11
26.	236-36	nil	One touching the same.	5;2
PAIR-COMPARISON TASK				
27.	47	SM *	They're the same but one's got white eyes and one's got black eyes. (Q) They're the same cos they've both got a fat head.	5;4

EXAMPLES OF INDIVIDUAL JUSTIFICATORY RESPONSES (cont.)

EXAMPLE	PAIR	SHAR-ING		JUSTIFICATION	C's AGE
28.	58	SM	*	They're the same but they're different, One's got not coloured eyes. But they're both standing-up eggs and they're both the same faces. (Q) They're just the same. (face = whole stimulus, not just M or S)	5;2
29.	15	SE	*	The mouths are not the same. (Q) The eyes are the same so the faces are the same.	4;4
30.	67		*	That's got sad and that's got smiling. (Q) They've got blue eyes the same (= black E).	5;9
31.	67			Same. No, not the same. You've got to look at them a lot!	6;6
32.	56	EM	*	They're the very same faces (p. E & M).	5;11
33.	28		*	They're the same because they're both happy. But there's one thing not the same: this one's an egg-shape (8) and this one's not an egg-shape (2). (Q) They're the same all over.	5;2
34.	28		*	Same. They've got the same eyes and the same faces. (face = Mouth)	5;7
35.	28		*	That one's got orange and that one's got orange eyes (= white E).	5;9
36.	34	EM	*	They're not quite the same. But they've both got straight mouths.	5;10
37.	45	nil		Not the same. One's a straight face and one's a smiling face. (face = Mouth)	4;11
38.	36			Not the same. One's got black eyes and a big head and he's smiley, and the other's got white eyes and a long head and he's sad.	5;5
39.	36			They're not at all the same (p. E & M).	5;11
40.	general comment			Sometimes the inside faces are the same but the outside heads have one going that way and one going that way (p. Shape orientations; Correct-matcher).	4;10
41.	"			If they're not the same shape and they've still got the same things on (p. M, E), they're not the same. (Correct-matcher)	5;4

"Unclassifiable" responses:

42.	56	EM		That one's the same (5) but that one isn't (6).	4;0

EXAMPLES OF INDIVIDUAL JUSTIFICATORY RESPONSES (cont.)

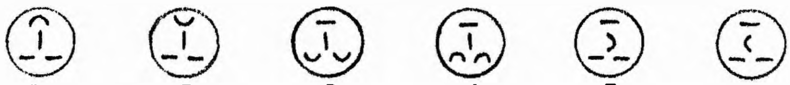
EXAMPLE	PAIR	SHARING		JUSTIFICATION	C's AGE
43.	56	EM	*	They're the same all over. (Refused further reason)	5;2
44.	34			That's the same (3) and that's not (4).	4;2
45.	16	E		They're not at all the same. (Refused further reason)	4;8
46.	45	nil		Not the same. But that's the same (re. 4. I.e., C had seen face 4 before, on a previous trial ??)	4;9

JUSTIFICATORY RESPONSES: EXPERIMENT 4, Conventional Pair-Comparison

EXAMPLE	PAIR		JUSTIFICATION	C's AGE
1.	12		That one's sad and that one's happy.	4;0
2.			It's a sad and a happy face.	4;5
3.			That one's happy and that one's sad.	4;9
4.			That one's sad and that one's smiling.	4;9
5.			No, cos that one's sad and that one's happy.	4;10
6.			That one's got a smiling face and that one's got a sad face.	5;0
7.	11		They're two smiling faces.	5;0
8.	22		(They're the same) Cos they're sad faces.	4;0
9.			Yes, they're both sad.	4;10
10.	34		The eyes are lines going that way and lines going that way (p. E, curve-up vs curve-down)	4;5
11.			(p. Eyes as not same)	4;8
12.			That one's eyes go down and that one's eyes go up.	4;8
13.			(p. Eyes as not same)	4;10
14.	56		The noses are one going that way and one going that way (p. N, curve-right vs curve-left).	4;5
15.			The noses are going not the same way.	4;9
16.		*	They're the same but the nose goes a different way. (Q) They're the same.	4;11
17.		*	Mouths the same.	4;11
18.		*	They're the same but the noses are going different ways. (Q) They're all the same faces.	5;0
19.			That one's got a bit that goes that way and that one's got a bit that goes that way (p. N, curve-right vs -left).	5;0
20.	66		They've got the same noses.	4;8
21.	15		No, but they're both happy.	4;10
22.			That one's got a happy face and the nose goes that way (p. M then N on 1, vs on 5).	4;11
23.	26		They're both sad.	4;10
24.	45		The noses have one going that way and one going that way (p. N, straight on 4, curve on 5).	4;5
25.	36		(p. Eyes as not same)	4;10
26.	55	*	(Not the same) They're funny. (Q) Because they're funny.	3;5
27.	46		They're not the same faces. (Q) (C refused further comment)	4;1

["unclassifiable" responses]

SAMPLE JUSTIFICATORY RESPONSES: EXPERIMENT 5, Inverted Stimuli



1 2 3 4 5 6

EXAMPLE	PAIR		JUSTIFICATION	C's AGE
28.	12		That one's got one that way and that one's got one that way (p. M, curve-up vs -down)	4;3
29.			Not the same. Well, they're <u>nearly</u> the same, but... (p. M as not same).	4;6
30.			The noses are... (p. M, then tried to turn cards right-way-up). That one faces that way and that one faces that way (p. M).	4;8
31.		*	(p. N as same)	4;10
32.	22		They're sad.	4;4
33.	34		Same, oh, not the same. One like that and one like that (p. E curve-up vs -down).	4;6
34.			The eyes (are not the same).	4;10
35.		*	(p. M & N as same).	4;10
36.	56	*	Not the same. No, they're the same. But that one's that way and that one's that way (p. N curve-left vs -right). (Q) Same.	4;3
37.			Same, no, different. They're <u>nearly</u> the same, but that's the other way (p. N).	4;6
38.			They're two... (p. N) They've got no "one". (<u>one</u> = straight N, i.e., resembling written number "1")	4;6
39.			That one's that way and that one's that way (p. N).	5;0
40.	55		They're sad.	4;4
41.	16		(p. M & N as not same)	4;3
42.	25		The mouth and the nose are different.	4;10
43.	13		The mouth is a line (p. M on 3). The nose is a line (p. N on 1 & 3). Those are... (p. E curve on 3 and straight on 1).	4;6
44.	23		It's a stroke mouth and a curly mouth. And that's got straight eyes (p. E on 2).	4;3
45.	24		The eyes are different.	4;1
46.	36		That one's curly and that one's straight (p. E).	4;3
47.	45		Two lines (p. straight E on 5 & N on 4), not the same (p. curved E on 4 & N on 5).	4;6
48.	46		Cos that one's got a nose that way (p. curved N on 6).	5;0

EXAMPLE	PAIR	JUSTIFICATION	C's AGE
		"Unclassifiable" responses:	
49.	55	* That one's that way (p. N).	4;3
50.	66	* That one's that way (p. N).	"
51.		* There's no "one" on it (i.e., it hasn't a straight N).	4;6
52.		* (p. N as not same)	4;6
53.	15	That one isn't the same (= 5) but that one is (= 1)	4;6
54.	16	That's a different face (= 6). (Q) That's the same (= 1) and that's not (= 6).	4;10
55.	35	That one's alright (= 3) and that one's not (= 5). (Q) (p. curved N on 5).	4;10

SAMPLE JUSTIFICATORY RESPONSES: EXPERIMENT 6

Pair-comparison with "pretraining"

1 2 3 4 5 6

EXAMPLE	PAIR		JUSTIFICATION	C's AGE
56.	12		Same. No, not the same. It's upside-down (p. M on 2).	3;9
57.			He's sad and he's happy	4;0
58.			The same eyes and the same nose. A sad mouth and a happy mouth.	4;6
59.			That one's this way and that one's this way (p. M)	4;9
60.	11		Cos they've got the mouths the same.	3;8
61.	34	*	Because they're the same faces. (Refused further comment)	3;9
62.			(p. E as not same)	4;2
63.			His eyes go up the way and his eyes go down the way.	4;5
64.			These eyes are this way and these eyes are this way (p. E).	4;9
65.	56		They go two different ways (p. N).	4;2
66.		*	It's straight eyes, a straight mouth, and the nose is curly. (Q re. N curl-left vs curl-right). They're the <u>same</u>.	4;7
67.			One's this way and one's that way (p. N)	4;9
68.			No, cos of the noses.	5;0
69.	55		(p. N as same)	3;9
70.	66		(p. N as same)	4;9
71.	23		It's a sad and a straight mouth.	4;2
72.	24		Same, not the same. Hasn't the same mouth.	3;8
73.	15	*	They're both happy and straight. (<u>happy</u> = M on both, <u>straight</u> = E on both)	4;0
74.	25		The mouth isn't the same, or not the nose.	3;8
75.	26		It's a straight mouth and a round mouth.	4;0
76.	35		The noses. (Q) They go that way (p. N).	4;9
77.	36		Same. Oh, no, not the same. His nose is round and his is straight.	4;0
78.	45		That one's nose goes like that (p. curved N on 5).	4;2
79.	46		The eyes aren't the same. But the mouths are the same.	4;6

EXPERIMENT 6: Samples of descriptions of stimuli given during prior "familiarisation" period

C's AGE	DESCRIPTION OF FACE 1 	DESCRIPTION OF FACE 2 
3;8 *	He's happy. (Q) He's asleep (re. E). (Q) [It's a straight thin nose.]	A funny mouth, he's crying. (Q)[He's awake.](Q) Got straight eyes. (Q)[Got a long thin nose.]
4;5	He's sleeping again (= straight E, seen already on another face). (Q) It's got a thin nose. (Q) The mouth is bent, it's up like that (p. M).	It's a sun (re. Shape circle). (Q) He's sour. (Q) He's sleeping (re. E). (Q) The nose is shut (= straight). (Q) The mouth is sour.
4;7	He's happy. (Q) He's sleeping, they're line eyes. (Q) It's a straight nose.	He's sad and he's sleeping. (Q) Straight nose.
4;10	It's a happy face. (Q) He's got a line nose and line eyes again.	He's got a sad mouth. (Q) He's got his eyes shut too. (Q) Got a line nose.

C's AGE	DESCRIPTION OF FACE 3 	DESCRIPTION OF FACE 4 
3;9 *	(p. E) The eyes look like an (?)oval, because they're squint. (Q) The mouth is straight and the nose is straight.	It's upside-down. (re. E). (Q) He's sleeping. (p. M & N:) Straight.
4;5	The eyes go down. (Q) The nose goes up the way. (Q) The mouth goes along the way.	The eyes are looking up. (Q) The nose is going up the way. (Q) The nose is going along the way and the mouth is going along the way.
4;9	A straight mouth. (Q) The eyes are this way (p. E). (Q) Up (re. E curve). (Q) That's a straight line (p. N).	The eyes are going down and the nose is just straight. (Q) The mouth is straight across.
5;0 *	It's a face with a nose like that, mouth like that, eyes like that (p. each in turn). They're right eyes, they're looking up. (Q) (p. N & M:) They're not bent.	He's looking down. (p. N & M:) They're straight.

KEY: * means the subject later gave a wrong "same" judgment of the face pair described on that line.

[ ] Phrases enclosed in square brackets were produced only after much prompting from the Experimenter.

EXPERIMENT 6: Samples of descriptions of stimuli given during prior "familiarisation" period (cont.)

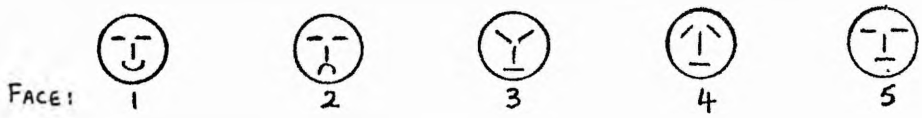
C's AGE	DESCRIPTION OF FACE 5 	DESCRIPTION OF FACE 6 
4;0 *	He's got the wrong nose. (Q re. N) That way (p. right for curve-right of N). (Q) The mouth and the eyes are straight.	He's got the wrong nose. (Q) Going that way (p. left for curve-left of N). (Q) He's got straight eyes and mouth.
4;6	It's an along the way face. (p. M). The eyes are straight. (Q) The nose is half-way round that way (p. N curve-right).	It's a bent nose. (Q) Goes like that (p. N, curve-left). (Q) It's straight eyes and straight mouth.
4;9 *	It's a face. (Q) He's asleep. (Q) The eyes are like "ones" (= straight lines, like written number "1"). The mouth is a "one". The nose is like a moon. (Q re. N) That way (p. N curve-right).	The nose is a moon, it's pointing back. The mouth and the eyes are lines.
4;9 *	Mouth that way, nose that way, eyes that way (p. each in turn). (Q) Got a square thin mouth. (Q) The eyes are closing. The mouth is closing. (Q) The nose is a rainbow. (Q) It goes touching that way (p. N curve-right).	The eyes is that way (p. E). (Q) Like a square window (re. E, = straight lines), and the mouth is that way as well. (Q) The nose is like a rainbow. (Q) It's that side way (p. N curve-left).

EXPERIMENT 6: Examples of complete profiles from preliminary "familiarisation" period descriptions

FACE (& ORDER of PRESENTATION)	DESCRIPTION
Subject aged 4;2 (entirely correct on subsequent pair-comparison)	
face 1 (1st)	It's a face. (Q) It's smiling. (Q) [The lines are straight lines.] (Q) [The nose is straight down.]
2 (5th)	The mouth is sad. (Q) The eyes are straight. (Q) The nose is straight.
3 (6th)	It's a sad face, sad eyes. (Q) The mouth is straight. (Q) The nose is straight.
4 (4th)	The mouth is straight and the eyes are bent, they're working up. (Q) A straight nose.
5 (2nd)	It's a funny face, I don't know that one. (Q) They're straight eyes. (Q) The nose is bent. (Q) The mouth is straight. (Q re. direction of N curve:) (p. right for N curve-right).
6 (3rd)	The nose is bent again. (Q) The eyes and the mouth are straight and the nose is bent like a sausage. (Q re. direction of N curve:) (p. left for N curve-left).
Subject aged 4;6 (entirely correct on subsequent pair-comparison)	
face 1 (5th)	He's happy. (Q) He's got long eyes. (Q) Got a long nose.
2 (3rd)	He's sad. (Q) A long nose. (Q) The eyes are long.
3 (4th)	The eyes are down-looking. (Q) It's a straight nose and mouth.
4 (6th)	(p. E:) They're round, they're looking up. (Q) The nose and the mouth are lines -- that one's down and that one's across.
5 (2nd)	It's a face. (Q) Got straight eyes and straight mouth. (Q) A round nose. (Q) Looking that way (p. N curve-right).
6 (1st)	It's a sun, a face. (Q) It's got a straight mouth, and straight eyes, but the nose is round. (Q) The nose is looking that way (p. N curve-left).

SAMPLE JUSTIFICATORY RESPONSES: EXPERIMENT 8,

pair-comparison with "affectively-distinct" stimuli



EXAMPLE	PAIR	JUSTIFICATION	C's AGE
1.	12	That mouth is down and that mouth is up,	3;2
2.		Happy, sad.	3;6
3.		(p. M) That's curving down and that's up.	3;9
4.		Sad and happy.	3;10
5.		He's got a round mouth, the nose is down, the eyes are straight (1); he's got a down mouth, a down nose, the eyes are straight (2).	4;2
6.		He's sad, he's happy.	4;8
7.	34	The eyes. His eyes are straight (3) and he's pulling on his (4). (C demonstrates, pulling own eyes down)	3;2
8.		He's got pulled-down eyes.(4).	3;4
9.		Sad (4) and happy (3).	3;6
10.		The eyes are pointed up and down.	3;9
11.		They're smiley faces (p. M) but his are pointing down and his are pointing up (p. E).	4;2
12.		He's miserable (4) and his eyes are up (3).	4;3
13.		The eyes, like you pull your eyes. He's pulling his eyes up and he's pulling his eyes down.	4;10
14.	15	Straight and curly (p. M).	3;9
15.		He's smiling and he's not smiling.	4;2
16.		He's smiling (1) and he's sad (5).	4;3
17.		He's sad (5) and he's happy (1).	4;4
18.		The mouths. One's round and up and one's straight.	5;2
19.	25	The mouth, it's down there (p. M on 2).	3;4
20.		He's got his mouth closed (5) and that one's not closed (2). And he's <u>sad</u>.(2).	3;11
21.		* The mouth is down, the faces are both sad.	3;11
22.		* The eyes are straight and they <u>might</u> be the same. (Q) Yes, they're the same cross faces.	4;2
23.		He's miserable (2) and he's happy (5).	4;3

SAMPLE JUSTIFICATORY RESPONSES: Experiment 8 (continued)

EXAMPLE	PAIR	JUSTIFICATION	C's AGE
24.	25	The mouths are not the same but the eyes are the same.	4;3
25.	35	He's cross (3).	3;2
26.		His eyes are up (3) and his are out (5).	3;11
27.		That one's got eyebrows going away (1) and that one's happy (5).	3;11
28.		(Same) He's happy (5). Oh, they're not the same; his eyes are up and his are down.	4;3
29.		(p. E as not same)	4;10
30.		The mouth is the same and the nose is the same but they're not the same eyes. His eyes are up (3) and his are down(5).	5;2
31.	45	His eyes are down (4).	3;2
32.		He's pushed his eyes up (4) and he's got his closed (5).	3;11
33.		That's an eye pointing down and that's an eye pointing straight.	4;2
34.	*	They're both sad.	4;3
35.		The same mouth and the same nose but not the same faces.	4;10
36.	13	The mouths are not the same. That One's happy (1) and that one's cross (3).	3;9
37.		The eyes are pointing up and straight.	3;9
38.		That one's closed his mouth (3) and that one's smiling (1).	3;11
39.		It's happy (1) and sad (3).	3;11
40.		His eyes are pointing up and his eyes are straight. That one's got a smile and that one's not smiley.	4;2
41.		That one's cross (3) and that one's happy (1).	4;3
42.	14	He's sad (4). The eyes are not the same, but the nose is the same.	3;2
43.		That one's mouth is closed and that one's smiling, and his eyes are up.	3;11
44.		The mouth is round, the nose is down, the eyes are straight (1); the mouth is straight, the nose is down and the eyes are down (4).	4;2

SAMPLE JUSTIFICATORY RESPONSES: Experiment 8 (continued)

EXAMPLE	PAIR		JUSTIFICATION	C's AGE
45.	14		(p. M & E) They're different.	4;3
46.			He's sad (4) and he's happy (1).	4;4
47.	23		Sad (2), happy (3).	3;6
48.			Straight and pointing up (p. E).	3;9
49.			One's sad and one's happy.	3;11
50.		*	(Not the same) That's a smiley face (3) and that's a cross face (2). (Hesitated) That's an eye pointing up (3). No, they're both the same because they're both cross.	4;2
51.			One eyes are straight and one eyes are up; a straight mouth and a curved mouth.	4;3
52.			That one's sad (2) and that one's cross (3).	4;3
53.	24		He's got a down mouth (2) and he's got pulled eyes (4).	3;2
54.			One's eyes is up, and that one's up (p. M on 2).	3;11
55.		*	He's a cross face (p. M on 2) and his eyes go down (4), his eyes are straight (2). They're the same because they're both cross faces.	4;2
56.			That one's cross (2) and that one's sad (4).	4;3
57.			The mouths and the eyebrows (p). That goes up and that one's straight (p. E).	5;2
58.	11		(p. E as same)	3;6
59.			They're both happy faces.	3;11
60.			That's straight (p. E) and there's a smile and the nose is straight.	4;2
61.	22		(p. M as same)	3;10
62.			Sad and sad.	4;9
63.	33		The eyes are pointing up. They've both got cross faces.	4;2
64.			The eyes are up and down, they're the same.	4;3
65.	44		That's a smile and the nose is straight and the eye is pointing down.	4;2
66.			The eyes and the mouths are the same.	4;3

SAMPLE JUSTIFICATORY RESPONSES: EXPERIMENT 8 (continued)

EXAMPLE	PAIR	JUSTIFICATION	C's AGE
67.	55	They're both sad.	4;4
68.		The eyes are the same and the nose is the same and the mouth is the same.	5;2

Unclassifiable / erroneous references

EXAMPLE	PAIR	JUSTIFICATION	C's AGE
69.	12	Broke the eyes off (p. M as not same). (E unclassifiable)	3;1
70.		That one's bad (p. M on 2) and that one's not bad (p. M on 1).	3;11
71.		* (p. M) It's like that.	4;8
72.	15	* (p. M as same)	4;4
73.	25	It's the same mouth.	4;8
74.	45	The mouths and the eyes are not the same. (M wrong)	3;9
75.		That's not the same bit (p. M) but the eyes are the same. (wrong -- it's M same, E not same)	4;3
76.	13	* The eyes are the same.	3;10
77.	11	* That's not the same (p. M).	4;3
78.	22	* Because they're just not the same <u>faces</u>.	3;11

Complete response records from two "affect-attending" subjects,
Experiment 8

1 2 3 4 5

TRIAL	PAIR	JUSTIFICATION	SUBJECT
1.	12	That one's happy and that one's sad.	girl
4.	34	That one's cross (3) and that one's sad (4).	age 3;9
7.	15	That one's sad (5) and that one's happy (1).	U-D
2.	25	That one's happy (5) and that one's sad (2).	alignment
12.	35	That One's cross (3) and that one's sad (5).	
15.	45	That One's happy (5) and that one's sad (4).	pair 12
10.	13	That one's happy and that one's cross.	first
8.	14	That One's happy and that one's sad.	
14.	23	That one's sad (2) and that one's cross (3).	
9.	24 *	That one's sad and that one's sad, they're the same.	
3.	11	They're happy.	
13.	22	That one's sad and that one's sad.	
5.	33	That one's cross and that one's cross.	
11.	44	---	
6.	55	---	

8.	12	(Same, no, not the same) Cross (2) and happy (1).	girl
1.	34	He's cross (4) and He's happy (3).	
14.	15	He's cross (5) and he's happy (1).	age 3;9
7.	25	Happy (5) and cross (2).	L-R
5.	35	He's happy (5) and he's cross (3).	alignment
13.	45	He's happy (5) and he's cross (4).	
10.	13	That's a happy face and that's a sad face.	pair 34
6.	14	Happy and cross.	first
9.	23	He's happy (3) and he's cross (2).	
3.	24	He's sad (4) and he's cross (2).	
2.	11	They're happy faces.	
4.	22	It's cross and cross.	
15.	33	They're two cross faces.	
11.	44	Both cross.	
12.	55	Happy and happy.	

REFERENCE NOTES

1. Clark, E.V. In Part 3 of the BBC Radio 4 series Babytalk, broadcast 1 June, 1982.

2. Wales, R.J. Unpublished data from an informally conducted "experiment" to demonstrate children's confusion errors in sameness judgments for a BBC film. Personal communication, University of St. Andrews, 1974.

3. Freeman, N. Preliminary survey of data on children's free drawings of faces. Personal communication, University of Bristol, 1978.

4. Heider, E.R. Rules for classification of encodings. Mimeo, Brown University, Providence, R.I. (Accompanying: Style and accuracy of verbal communication within and between social classes. Doctoral dissertation, Harvard University, 1968.)

5. Chalmers, M. Personal communication, University of Edinburgh, 1980.

REFERENCES

Ahrens, R. Beitrag zur Entwicklung des Physiognomie -- und Mimikerkennens. Zeitschrift für experimentelle und angewandte Psychologie, 1954, 2, 412-454, 599-633.

Al-Issa, I. The development of word definition in children. Journal of Genetic Psychology, 1969, 114, 25-28.

Baron, J. Intelligence and general strategies. In G. Underwood (Ed.), Strategies of Information Processing. London: Academic Press, 1978, pp. 403-450.

Barroso, F. & Braine, L.G. "Mirror-image" errors without mirror-image stimuli. Journal of Experimental Child Psychology, 1974, 18, 213-225.

Bassili, J.N. Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. Journal of Personality and Social Psychology, 1979, 37, 2049-2058.

Berthoud, M. & Vurpillot, E. Influence du nombre de différences sur les réponses "pas pareil" chez l'enfant d'âge préscolaire. Enfance, 1970, 1, 23-30.

Bolton, N. The Psychology of Thinking. London: Methuen, 1972.

Bornstein, M.H. & Gross, C.G. Perceptual similarity of mirror images in infancy. Cognition, 1978, 6, 89-116.

Bourne, L.E. & Restle, F. Mathematical theory of concept identification. Psychological Review, 1959, 66, 278-296.

Bozinou, E. & Goulet, L.R. Acquisition and transfer of sorting mechanisms in discrimination learning and free recall by nursery school children. Child Development, 1974, 45, 816-820.

Braine, L.G. Age changes in mode of perceiving geometric forms. Psychonomic Science, 1965, 2, 155-156.

Brooks, R.M. & Goldstein, A.G. Recognition by children of inverted photographs of faces. Child Development, 1963, 34, 1033-1040.

Brown, D.R. & Goldstein, J.A. Stimulus correlates of the discrimination behaviour of children. Psychonomic Science, 1967, 9, 177-178.

Bruner, J.S., Goodnow, J.J. & Austin, G.A. A Study of Thinking. New York: Wiley, 1956.

Bruner, J.S., Olver, R.R. & Greenfield, P.M. (Eds.). Studies in Cognitive Growth. New York: Wiley, 1966.

Brunswik, E. Perception and the Representative Design of Psychological Experiments. Berkeley & Los Angeles: University of California Press, 1956.

Brunswik, E. & Reiter, L. Eindruckscharaktere schematisierter Gesichter. Zeitschrift für Psychologie, 1938, 142, 67-134.

Bryant, P.E. Perception and Understanding in Young Children: An Experimental Approach. London: Methuen, 1974.

Campbell, R., Donaldson, M. & Young, B. Constraints on classificatory skills in young children. British Journal of Psychology, 1976, 67, 89-100.

Campbell, R. & Wales, R.J. The study of language acquisition. In J. Lyons (Ed.), New Horizons in Linguistics, Vol. I. London: Penguin, 1973 edition.

Carey, S. & Diamond, R. From piecemeal to configural representation of faces. Science, 1977, 195, 312-314.

Chalmers, M. Why are young children's verbal descriptions inadequate? Paper presented at the University of St. Andrews, Department of Psychology, 31 October, 1980.

Claparède, E. La conscience de la ressemblance et de la différence chez l'enfant. Archives de Psychologie, 1919, 17, 67-80.

Clark, E.V. On the acquisition of the meaning of "before" and "after". Journal of Verbal Learning and Verbal Behaviour, 1971, 10, 266-275.

Clark, E.V. What's in a word? On the child's acquisition of semantics in his first language. In T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language. London & New York: Academic Press, 1973, pp. 65-110. (a)

Clark, E.V. Non-linguistic strategies and the acquisition of word meanings. Cognition, 1973, 2, 161-182. (b)

Clark, H.H. The primitive nature of children's relational concepts. In J.R. Hayes (Ed.), Cognition and the Development of Language. New York: Wiley, 1970.

Clark, H.H. Time, space, semantics and the child. In T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language. London & New York: Academic Press, 1973, pp. 28-64.

Clark, H.H. & Clark, E.V. Psychology and Language: An Introduction to Psycholinguistics. New York (etc.): Harcourt Brace Jovanovich, 1977.

Clark, H.J. Recognition memory for random shapes as a function of complexity, association value, and delay. Journal of Experimental Psychology, 1965, 69, 590-595.

Corah, N.L. The influence of some stimulus characteristics on color and form perception in nursery-school children. Child Development, 1966, 37, 205-211.

Danks, J.H. Encoding of novel figures for communication and memory. Cognitive Psychology, 1970, 1, 179-191.

Diamond, R. & Carey, S. Developmental changes in the representation of faces. Journal of Experimental Child Psychology, 1977, 23, 1-22.

Donaldson, M. Children's Minds. Glasgow: Fontana, 1978.

Donaldson, M. & Balfour, G. Less is more: A study of language comprehension in children. British Journal of Psychology, 1968, 59, 461-471.

Donaldson, M. & Wales, R.J. On the acquisition of some relational terms. In J.R. Hayes (Ed.), Cognition and the Development of Language. New York: Wiley, 1970, pp. 235-268.

Elkind, D. Ambiguous pictures for study of perceptual development and learning. Child Development, 1964, 35, 1391-1396.

Elkind, D. Koegler, R.R. & Go, E. Studies in perceptual development, II: part-whole perception. Child Development, 1964, 35, 81-90.

Elkind, D. & Weiss, J. Studies in perceptual development, III: perceptual exploration. Child Development, 1967, 38, 553-561.

Fantz, R.L. Pattern vision in young infants. Psychological Record, 1958, 8, 43-47.

Fantz, R.L. The origin of form perception. Scientific American, 1961, 204, 66-72.

Fantz, R.L. Pattern vision in newborn infants. Science, 1963, 140, 296-297.

Fantz, R.L. Pattern discrimination and selective attention as determinants of perceptual development from birth. In A.H. Kidd & J.L. Rivoire (Eds.), Perceptual Development in Children. New York: International University Press, 1966, pp. 143-173.

Farkas, M. & Elkind, D. Effects of distance and stimulus position on perceptual comparison at five age levels. Child Development, 1974, 45, 184-188.

Faw, T.T. & Nunally, J.C. The influence of stimulus complexity, novelty, and affective value on children's visual fixations. Journal of Experimental Child Psychology, 1968, 6, 141-153.

Flavell, J.H. (with others). The Development of Role-Taking and Communication Skills in Children. New York: Robert Krieger, 1975 edition.

Frijda, N.H. Recognition of emotion. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, Vol. 4. London & New York: Academic Press, 1969, pp. 167-223.

Garner, W.R. Uncertainty and Structure as Psychological Concepts. New York: Wiley, 1962.

Garner, W.R. To perceive is to know. American Psychologist, 1966, 21, 11-19.

Garner, W.R. The stimulus in information processing. American Psychologist, 1970, 25, 350-358.

Garner, W.R. The Processing of Information and Structure. Potomac, Md.: Laurence Erlbaum, 1974.

Garner, W.R. & Clement, D.E. Goodness of pattern and pattern uncertainty. Journal of Verbal Learning and Verbal Behaviour, 1963, 2, 446-452.

Garner, W.R. & Felfoldy, G.L. Integrality of stimulus dimensions in various types of information processing. Cognitive Psychology, 1970, 1, 225-241.

Gathercole, V.C. On the production of comparative structures in child speech. Paper presented at the Summer Meeting of the Linguistic Society of America, July 28, 1978.

Gati, I. & Tversky, A. Representations of qualitative and quantitative dimensions. Journal of Experimental Psychology: Human Perception and Performance, 1982, 8, 325-340.

Gelman, R. & Gallistel, C.R. The Child's Understanding of Number. Cambridge, Mass.: Harvard University Press, 1978.

Ghent, L. Recognition by children of realistic figures in various orientations. Canadian Journal of Psychology, 1960, 14, 249-256.

Ghent, L. Form and its orientation; a child's eye view. American Journal of Psychology, 1961, 74, 177-190.

Ghent, L. & Bernstein, L. Influence of the orientation of geometric forms on their recognition by children. Perceptual and Motor Skills, 1961, 12, 95-101.

Ghent, L., Bernstein, L. & Goldweber, A.M. Preferences for orientation of form under varying conditions. Perceptual and Motor Skills, 1960, 11, 46.

Gibson, E.J. Development of perception: Discrimination of depth compared with discrimination of graphic symbols. In J.C. Wright & J. Kagan (Eds.), Basic Cognitive Processes in Children. Monographs of the Society for Research in Child Development, Serial No. 86, 1963, 28(2), pp. 5-32.

Gibson, E.J. Learning to read. Science, 1965, 148, 1066-1072.

Gibson, E.J. Principles of Perceptual Learning and Development. New York: Appleton-Century-Crofts, 1969.

Gibson, E.J. Perceptual learning and the theory of word perception. Cognitive Psychology, 1971, 2, 351-368.

Gibson, E.J., Gibson, J.J., Pick, A.D. & Osser, H. A developmental study of the discrimination of letter-like forms. Journal of Comparative and Physiological Psychology, 1962, 55, 897-906.

Gibson, J.J. & Robinson, D. Orientation in visual perception; The recognition of familiar plane forms in differing orientations. Psychological Monographs, 1935, 46, 39-47.

Glucksberg, S., Hay, A. & Danks, J.H. Words in utterance contexts: Young children do not confuse the meanings of same and different. Child Development, 1976, 47, 737-741.

Glucksberg, S. & Krauss, R.M. What do people say after they have learned how to talk? Studies of the development of referential communication. Merrill-Palmer Quarterly of Behaviour and Development, 1967, 13, 309-316.

Glucksberg, S., Krauss, R.M. & Weisberg, R. Referential communication in nursery school children: method and some preliminary findings. Journal of Experimental Child Psychology, 1966, 3, 333-342.

Goldstein, A.G. & Chance, J.E. Recognition of children's faces. Child Development, 1964, 35, 129-136.

Goldstein, A.G. & Mackenberg, E. Recognition of human faces from isolated facial features. Psychonomic Science, 1966, 6, 149-150.

Goodman, N. Seven strictures on similarity. In L. Foster & J.W. Swanson (Eds.), Experience and Theory. London: Duckworth, 1970, pp. 19=29.

Goodnow, J. Rules and repertoires, rituals and tricks of the trade: Social and informational aspects to cognitive representational development. In S. Farnham-Diggory (Ed.), Information Processing in Children. New York: Academic Press, 1972.

Gottschalk, J., Bryden, M.P. & Rabinovitch, S. Spatial organization of children's responses to a pictorial display. Child Development, 1964, 35, 811-815.

Gregson, R.A.M. Psychometrics of Similarity. New York & London: Academic Press, 1975.

Grieve, R. & Hoogenraad, R. Using language if you don't have much. In R.J. Wales & E.C.T. Walker (Eds.), New Approaches to Mechanisms in Language. Amsterdam: North-Holland, 1976, pp. 1-28.

Hale, G.A. & Lipps, L.E.T. Stimulus matching and component selection: Alternative approaches to measuring children's attention to stimulus components. Child Development, 1974, 45, 383-388.

Halford, G.S. & Macdonald, C. Children's pattern construction as a function of age and complexity. Child Development, 1977, 48, 1096-1100.

Heider, E.R., Cazden, C.B. & Brown, R. Social class differences in the effectiveness and style of children's coding ability. Project Literacy Report, No. 9, Cornell University, Ithaca, N.Y., 1968.

Hendrickson, L. & Muehl, S. The effect of attention and motor response pretraining on learning to discriminate B and D in kindergarten children. Journal of Educational Psychology, 1962, 53, 236-241.

Hess, V.L. & Pick, A.D. Discrimination of schematic faces by nursery school children. Child Development, 1974, 45, 1151-1154.

Hunton, V. The recognition of inverted pictures by children. Journal of Genetic Psychology, 1955, 86, 281-288.

Huttenlocher, J. Development of formal reasoning on concept formation problems. Child Development, 1964, 35, 1233-1242.

Huttenlocher, J. Discrimination of figure orientation: effects of relative position. Journal of Comparative and Physiological Psychology, 1967, 63, 359-361. (a)

Huttenlocher, J. Children's ability to order and orient objects. Child Development, 1967, 38, 1169-1176. (b)

Imai, S. & Garner, W.R. Discriminability and preference for attributes in free and constrained classification. Journal of Experimental Psychology, 1965, 69, 596-608.

Inhelder, B. & Piaget, J. The Early Growth of Logical Thinking. New York: Norton, 1964.

Jeffrey, W.E. Variables in early discrimination learning: I. Motor responses in the training of a left-right discrimination. Child Development, 1958, 29, 269-275.

Jeffrey, W.E. Discrimination of oblique lines by children. Journal of Comparative and Physiological Psychology, 1966, 62, 154-156.

Johnston, R.P. & Singleton, C.H. Social class and communication style: The ability of middle and working class five year olds to encode and decode abstract stimuli. British Journal of Psychology, 1977, 68, 237-244.

Jörg, S. & Hörmann, H. The influence of general and specific verbal labels on the recognition of labelled and unlabelled parts of pictures. Mimeo, Psychologisches Institut, Ruhr-Universität Bochum, West Germany, 1977.

Kagan, J., Moss, H.A. & Sigel, I.E. Psychological significance of styles of conceptualization. In J.C. Wright & J. Kagan (Eds.), Basic Cognitive Processes in Children. Monographs of the Society for Research in Child Development, Serial No. 86, 1963, 28(2), pp. 73-124.

Karmiloff-Smith, A. More about the same: children's understanding of post-articles. Journal of Child Language, 1977, 4, 377-394.

Kemler, D.G. & Smith, L.B. Is there a developmental trend from integrality to separability in perception? Journal of Experimental Child Psychology, 1978, 26, 498-507.

Kemler, D.G. & Smith, L.B. Accessing similarity and dimensional relations: The effects of integrality and separability on the discovery of complex concepts. Journal of Experimental Psychology: General, 1979, 108, 133-150.

Kendler, H.H. & Kendler, T.S. Vertical and horizontal processes in problem solving. Psychological Review, 1962, 62, 1-16.

Kogan, N. Cognitive Styles in Infancy and Early Childhood. Hillsdale, N.J.: Laurence Erlbaum, 1976.

Krumhansl, C.L. Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. Psychological Review, 1978, 85, 445-463.

Langdell, T. Recognition of faces: An approach to the study of autism. Journal of Child Psychology and Psychiatry, 1978, 19, 255-268.

Lasky, R.E. The ability of six-year olds, eight-year olds, and adults to abstract visual patterns. Child Development, 1974, 45, 626-632.

Levelt, W.J.M. Hierarchical clustering algorithms in the psychology of grammar. In G.B. Flores d'Arcais & W.J.M. Levelt (Eds.), Advances in Psycholinguistics. Amsterdam: North-Holland, 1970, Ch. 4.

Lian, A. The Psychological Study of Object Perception: Examination of Methodological Problems and a Critique of Main Research Approaches. London & New York: Academic Press, 1981.

Long, L. Size discrimination in children. Child Development, 1941, 12, 247-254.

Mackworth, N.H. & Bruner, J.S. How adults and children search and recognize pictures. Human Development, 1970, 13, 149-177.

Macrae, A.J. Meaning relations in language development: A study of some converse pairs and directional opposites. Doctoral dissertation, University of Edinburgh, 1976.

Macrae, A.J. Natural descriptions and models of semantic development. Paper presented to the Child Language Seminar, University of York, April, 1978.

McConnell, O.L. Perceptual versus verbal mediation in the concept learning of children. Child Development, 1964, 35, 1373-1383.

McGurk, H. The salience of orientation in young children's perception of form. Child Development, 1972, 43, 1047-1052.

Mwanalushi, M. Imaginal factors in the coding of random patterns by children. Child Development, 1974, 45, 204-207.

Nelson, K. The relation of form recognition to concept development. Child Development, 1972, 43, 67-74.

Nelson, K. Some evidence for the cognitive primacy of categorization and its functional basis. Merrill-Palmer Quarterly of Behavior and Development, 1973, 19, 21-39.

Nelson, K. Concept, word and sentence: Interrelations in acquisition and development. Psychological Review, 1974, 81, 267-285. (a)

Nelson, K. Variations in children's concepts by age and category. Child Development, 1974, 45, 577-584. (b)

Nelson, K. Semantic development and the development of semantic memory. In K.E. Nelson (Ed.), Children's Language, Vol. 1. New York: Gardner Press, 1978, pp. 39-80. (a)

Nelson, K. Developing a functional semantic system. Paper presented at the Child Language Seminar, University of York, April, 1978. (b)

Odom, R.D. & Lemond, C.M. Children's use of component patterns of faces in multidimensional recall problems. Child Development, 1974, 45, 527-531.

Odom, R.D. & Mumbauer, C.C. Dimensional salience and identification of the relevant dimension in problem solving. Developmental Psychology, 1971, 4, 135-140.

Oléron, P. Le développement des réponses à la relation identité-dissemblance. Ses rapports avec le langage. Psychologie Français, 1962, 7, 4-16.

Olson, D.R. Cognitive Development: The Child's Acquisition of Diagonality. New York: Academic Press, 1970.

Olson, D.R. & Baker, N.E. Children's recall of spatial orientation of objects. Journal of Genetic Psychology, 1969, 114, 273-281.

Osler, S.F. & Kofsky, E. Stimulus uncertainty as a variable in the development of conceptual ability. Journal of Experimental Child Psychology, 1965, 2, 264-279.

Over, R. & Over, J. Detection and recognition of mirror-image obliques by young children. Journal of Comparative and Physiological Psychology, 1967, 64, 467-470.

Pick, A.D. Improvement of visual and tactual form discrimination. Journal of Experimental Psychology, 1965, 69, 331-339.

Pick, A.D. & Frankel, G.W. A developmental study of strategies of visual selectivity. Child Development, 1974, 45, 1162-1165.

Pick, A.D., Hales, J.J., Christy, M.D., Frankel, G.W. & Glick, J. The effect of a human facial context on the discrimination and recognition of curved lines. Psychonomic Science, 1972, 27, 239-242.

Posnansky, C.J. & Neumann, P.G. The abstraction of visual prototypes by children. Journal of Experimental Child Psychology, 1976, 21, 367-379.

Posner, M.I. Abstraction and the process of recognition. In G.H. Bower & J.T. Spence (Eds.), The Psychology of Learning and Motivation, Vol. 3. New York: Academic Press, 1969.

Reed, S.K. Pattern recognition and categorization. Cognitive Psychology, 1972, 3, 382-407.

Reichard, S., Schneider, M. & Rapaport, D. The development of concept formation in children. American Journal of Orthopsychiatry, 1944, 14, 156-161.

Ricciuti, H. Object-grouping and selective ordering behavior in infants 12-24 months old. Merrill-Palmer Quarterly of Behavior and Development, 1965, 11, 129-148.

Rice, C. The orientation of plane figures as a factor in their perception by children. Child Development, 1930, 1, 111-143.

Richards, M. Sorting out what's in a word from what's not. Evaluating Clark's semantic features acquisition theory. Journal of Experimental Child Psychology, 1979, 27, 1-47.

Rock, I. Halper, F. & Clayton, T. The perception and recognition of complex figures. Cognitive Psychology, 1972, 3, 655-673.

Rosch, E.H. On the internal structure of perceptual and semantic categories. In T.E. Moore (Ed.), Cognitive Development and the Acquisition of Language. London & New York: Academic Press, 1973. (a)

Rosch, E.H. Natural categories. Cognitive Psychology, 1973, 4, 328-350. (b)

Rosch, E. Cognitive reference points. Cognitive Psychology, 1975, 7, 532-547.

Rosch, E. Human categorization. In N. Warren (Ed.), Advances in Cross-cultural Psychology, Vol. 1. London: Academic Press, 1977.

Rosch, E. & Mervis, C.B. Family resemblances: Studies in the internal structure of cateogires. Cognitive Psychology, 1975, 7, 573-605.

Rosch, E., Mervis, C.B., Gray, W., Johnson, D. & Boyes-Braem, P. Basic objects in natural categories. Cognitive Psychology, 1976, 8, 382-439.

Rosenblith, J.F. Judgments of simple geometric forms by children. Perceptual and Motor Skills, 1965, 21, 947-990.

Rossi, E. Development of classificatory behavior. Child Development, 1964, 35, 137-142.

Rudel, R.G. & Teuber, H.L. Discrimination of direction of line in children. Journal of Comparative and Physiological Psychology, 1963, 56, 892-898.

Saltz, E. & Finkelstein, C. Does imagery retard conceptual behavior? Child Development, 1974, 45, 1093-1097.

Saltz, E. & Sigel, I.E. Concept overdiscrimination in children. Journal of Experimental Psychology, 1967, 73, 1-8.

Saltz, E., Soller, E. & Sigel, I.E. The development of natural language concepts. Child Development, 1972, 43, 1191-1202.

Santa Barbara, J.F. Jnr. & Paré, W.P. Information processing in pre-school children. Psychonomic Science, 1965, 2, 143-144.

Santostefano, S. & Paley, E. Development of cognitive controls in children. Child Development, 1964, 35, 939-949.

Schaller, M.J. & Harris, L.J. Children judge "perspective" transforma-ations of letterlike forms as different from prototypes. Journal of Experimental Child Psychology, 1974, 18, 226-241.

Schwarz, G. & Tversky, A. On the reciprocity of proximity relations. Journal of Mathematical Psychology, 1980, 22, 157-175.

Sekuler, R.W. & Rosenblith, J.F. Discrimination of direction of line and the effect of stimulus alignment. Psychonomic Science, 1964, 1, 143-144.

Seymour, P.H.K. Asymmetries in judgements of verticality. Journal of Experimental Psychology, 1974, 102, 447-455.

Seymour, P.H.K. Human Visual Cognition: A Study in Experimental Cognitive Psychology. London: Collier Macmillan, 1979.

Shepard, R.N. Attention and the metric structure of the stimulus space. Journal of Mathematical Psychology, 1964, 1, 54-87.

Shepard, R.N. & Chipman, S. Second-order isomorphism of internal representations: Shapes of states. Cognitive Psychology, 1970, 1, 1-17.

Siegel, S. Nonparametric Statistics for the Behavioural Sciences. New York: McGraw-Hill, 1956.

Sigel, I.E. Developmental trends in the abstraction ability of children. Child Development, 1953, 24, 131-144.

Sinclair, H. & Piaget, J. Sondage sur l'identité. Les formes d'équivalence et la conservation lors de la rotation d'un carré. In J. Piaget, H. Sinclair & Vinh-Bang (Eds.), Epistémologie et psychologie de l'identité, Paris: Presses Universitaires de France, 1968, pp. 123-144.

Sjöberg, L. A study of four methods for scaling paired comparisons data. Reports from the Psychological Laboratories, The University of Stockholm, No. 196, July 1965.

Smith, L.B. & Kemler, D.G. Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. Journal of Experimental Child Psychology, 1977, 24, 279-298.

Smith, L.B. & Kemler, D.G. Levels of experienced dimensionality in children and adults. Cognitive Psychology, 1978, 10, 502-532.

Steinberg, B.M. Information processing in the third year: Coding, memory, transfer. Child Development, 1974, 45, 503-507.

Strayer, J. & Ames, E.W. Stimulus orientation and the apparent developmental lag between perception and performance. Child Development, 1972, 43, 1345-1354.

Suchman, R.G. & Trabasso, T. Stimulus preference and cue function in young children's concept attainment. Journal of Experimental Child Psychology, 1966, 3, 188-198.

Suppes, P. & Rosenthal-Hill, I. Concept formation by kindergarten children in a card-sorting task. Journal of Experimental Child Psychology, 1968, 6, 212-230.

Taylor, J.A. Form discrimination in young children and the concept of similarity. Doctoral dissertation, University of Edinburgh, 1973.

Taylor, J.A. & Wales, R.J. A developmental study of form discrimination in pre-school children. Quarterly Journal of Experimental Psychology, 1970, 22, 720-734.

Torgerson, W.S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Treisman, A.M. & Gelade, G. A feature-integration theory of attention. Cognitive Psychology, 1980, 12, 97-136.

Tversky, A. Features of similarity. Psychological Review, 1977, 84, 327-352.

Tversky, A. & Krantz, D. Similarity of schematic faces: a test of interdimensional additivity. Perception and Psychophysics, 1969, 5, 124-128.

Vanderplas, J.M. & Garvin, E.A. Complexity, association value, and practice as factors in shape recognition following paired-associates training. Journal of Experimental Psychology, 1959, 57, 155-163.

Vernon, M.D. Perception in relation to cognition. In A.H. Kidd & J.L. Rivoire (Eds.), Perceptual Development in Children. New York: International Universities Press, 1966, pp. 391-406.

Vurpillot, E. The development of scanning strategies and their relation to visual differentiation. Journal of Experimental Child Psychology, 1968, 6, 632-650.

Vurpillot, E. Influence de la nature d'une différence sur la détection par des enfants d'âge préscolaire. Enfance, 1969, 22, 149-164.

Vurpillot, E. The Visual World of the Child. London: Allen & Unwin, 1976.

Vurpillot, E., Lecuyer, R., Moal, A. & Pinean, A. Perception de déplacements et jugements d'identité chez l'enfant d'âge préscolaire. Année Psychologique, 1971, 71, 31-52.

Vurpillot, E. & Moal, A. Evolution des critères d'identité chez des enfants d'âge préscolaire dans une tâche de différentiation perceptive. Année Psychologique, 1970, 70, 391-406.

Vygotsky, L.S. Thought and Language. Cambridge, Mass.: M.I.T. Press, 1962.

Wales, R.J. & Campbell, R.N. On the development of comparison and the comparison of development. In G.B. Flores d'Arcais & W.J.M. Levelt (Eds.), Advances in Psycholinguistics. Amsterdam: North-Holland, 1970, pp. 373-396.

Wales, R.J., Garman, M.A.G. & Griffiths, P.D. More or less the same: a markedly different view of children's comparative judgements in three cultures. In R.J. Wales & E.C.T. Walker (Eds.), New Approaches to Mechanisms in Language. Amsterdam: North-Holland, 1976.

Watson, J.S. Perception of object orientation in infants. Merrill-Palmer Quarterly of Behavior and Development, 1966, 12, 73-94.

Webb, R.A., Oliveri, M.E. & O'Keeffe, L. Investigations of the meaning of "different" in the language of young children. Child Development, 1974, 45, 984-991.

Whitman, J.R. & Garner, W.R. Concept learning as a function of form of internal structure. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 195-202.

Williams, T.M., Aiken, L.S. & Gordon, R. Developmental reversal in classification accuracy due to an increase with age in orientation-free judgments. Journal of Experimental Child Psychology, 1976, 21, 128-143.

Winer, B.J. Statistical Principles in Experimental Design. London: McGraw-Hill Kogakusha, 1971 edition.

Wohlwill, J.F. & Wiener, M. Discrimination of form orientation in young children. Child Development, 1964, 35, 1113-1125.

Zaporozhets, A.V. The development of perception in the preschool child. In P.H. Mussen (Ed.), European Research in Cognitive Development. Monographs of the Society for Research in Child Development, 1965, 30(2), pp. 82-101.

Zusne, L. Moments of area and of the perimeter of visual form as predictors of discrimination performance. Journal of Experimental Psychology, 1965, 69, 213-220.

Zusne, L. Visual Perception of Form. New York & London: Academic Press, 1970.