Sexing Day-Old Chicks: A Case Study and Expert Systems Analysis of a Difficult Perceptual-Learning Task

Irving Biederman
State University of New York at Buffalo

Margaret M. Shiffrar
University of California at Santa Cruz

The sexing of day-old chicks has been regarded as an extraordinarily difficult perceptual task requiring years of extensive practice for its mastery. Experts can sex chicks at over 98% accuracy at a rate of 1,000 chicks per hour spending less than a half second viewing the cloacal region. Naive subjects were shown 18 pictures of cloacal regions of male and female chicks (in random appearing arrangement) and asked to judge the sex of each chick. The pictures included a number of rare and difficult configurations. The subjects were then instructed as to the location of a critical cloacal structure for which a simple contrast in shape (convex vs. concave or flat) could serve as an indicant of sex. When the subjects judged the pictures again (in a different order), accuracy increased from slightly above chance to a level comparable to that achieved by a sample of experts. The correlation (over items) between the naive subjects and the experts before instruction was .21; after instruction, .82. The instructions were based on an interview and observation of an expert who had spent 50 years sexing 55 million chicks. Much of the reported difficulty in developing perceptual expertise in this task may stem from the need to classify extremely rare configurations in which the convexity of the structure is not apparent. The rate of learning of these instances could be greatly increased through the use of simple instructions that specified the location of diagnostic contour contrasts. A parallel is drawn between learning to sex chicks and learning to classify tanks as friend or foe.

This investigation was directed toward an understanding of the perceptual learning required for the sexing of day-old chicks. It has been reported that years are spent acquiring this skill, an absolutely critical skill for the success of modern hatcheries (Lunn, 1948). Professional sexers can classify 1,000 chicks per hour at over 98% accuracy. Less than a second is actually spent looking at each chick. Given that such high performance levels are possible, why should the perceptual learning required for this activity be so difficult? What is it that is learned? Under what conditions will any visual task that can be ultimately performed at a high level of speed and accuracy be difficult to learn?

Background

Like many perceptual psychologists, the first author (Irving Biederman) was introduced to this problem by the first figure in E. J. Gibson's *Principles of Perceptual Learning and Development* (1969). Professor Gibson had reproduced a figure from Canfield (1941) that showed the genital eminences of pullets and cockerels (Figure 1). Despite a considerable initial effort

and occasional attempts distributed over almost two decades at studying this figure, Biederman was unable to determine what information could be used to distinguish males from females. His experience was not unique: A number of other perceptual and cognitive psychologists that we questioned had also studied that figure and all but one (not verified before the results of this research were discussed with him) admitted to being unable to derive a perceptual basis for the discrimination.

For commercial egg producers, it is necessary to sort male chicks from female chicks as early as possible before considerable feed costs are invested in a useless bird. Males also deny the females easy access to food and water and, in general, cause sufficient turmoil in the hatcheries to significantly reduce the productivity of the pullets. Secondary sex characteristics, such as the rooster's comb, only appear a month after hatching.

Despite the economic interest in telling male from female chicks over the centuries of the domestication of this bird, it was only in the late 1920s that a perceptual basis for this discrimination was discovered by the Japanese. A Japanese delegation visiting the University of British Columbia in 1934 demonstrated their methods to representatives from hatcheries in the state of Washington. So important is this discrimination to the commercial success of hatcheries that in that same year schools whose sole purpose was the training of sexers were founded in Washington and Petaluma, California (Lunn, 1948).

The opportunity for an expert systems approach was suggested by an article from the October 22, 1984 *San Francisco Chronicle* (Carroll, 1984) entitled "A Lifetime Spent Looking at Chicks." This piece described the professional life of Heimer Carlson of Petaluma, who had spent 50 years typing 55 million chicks. Mr. Carlson had been in the very first class at the Petaluma school.

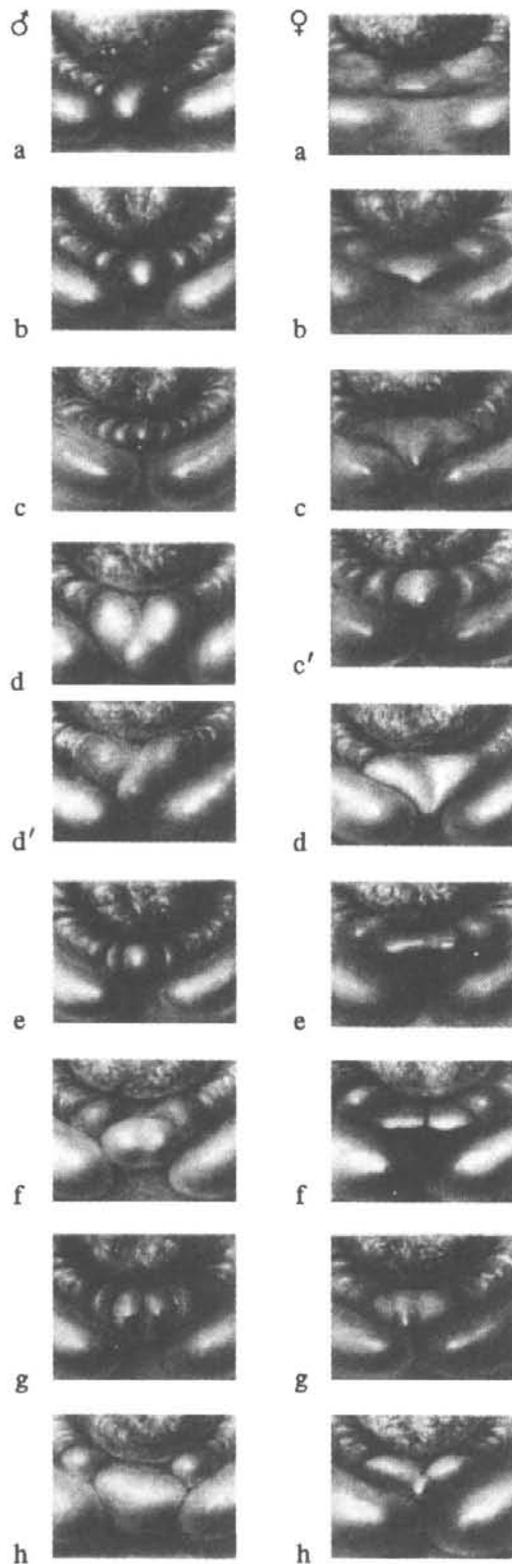Mr. Carlson, who was in semiretirement, allowed us to ob-

Figure 1. Genital eminences of day-old chicks. (From "Sex Determination of Day-Old Chicks II: Type Variations" by T. H. Canfield, 1941, *Poultry Science, 20,* pp. 327–328. Copyright 1941 by the Poultry Science Association, Inc. Reprinted by permission. Male genitals are in the left column, female genitals in the right colmn.)

serve and interview him while he sexed several hundred chickens at the hatchery at the University of California at Davis.

Observations

The chicks, only a few hours old, are brought to the sexer in trays of 100. The task requires that the cloaca be everted. The chick is held in the left hand (for a right-handed person) and the fecal contents are squirted into a container to clear the cloaca (see Figure 2). Gentle but firm pressure from the two thumbs and right forefinger are exerted to spread the ventral surface of the cloaca upwards to expose the eminence, called the "bead." The eminence is about the size of a pin head. The sexing decision must be made quickly because the chick is at risk from the vent eversion. Females are traditionally placed in a tray on the right and males on the left.

Mr. Carlson worked quickly and steadily, spending approximately .5 s actually looking at each eminence under magnification and a bright 200-W bulb.

Expert Systems Analysis

Because of the similarities among the eminences (Figure 1) it is clear that a naive observer would not know where to look in the picture or what to look for. We asked Mr. Carlson to circle the critical areas in each of the pictures on a copy of Figure 1. The critical region, the bead, is located between or above the opening in the transverse folds (the lower symmetrical lobes) or the longitudinal folds (the upper symmetrical processes lying on the transverse folds).

Our examination of the beads revealed a simple difference in the contours between males and females. In males, the eminence was convex; in females, flat or concave. This differentiation corresponded to the descriptions offered by some of the sexers who described males as "round" and females as "pointy." When a sketch of the shapes of the eminences was drawn (similar to those in Figure 3) and shown to Mr. Carlson, he agreed as to its diagnosticity.

This contrast in contour (concave vs. convex) has the characteristic that it is invariant over viewpoint. That is, from virtually any viewpoint of an object, convex regions will remain convex, concave regions will remain concave. Thus, a two-dimensional projection of a convex or concave section of contour to the eye will, with extremely high likelihood, be produced by a convex or concave contour, respectively, in the three-dimensional world. Lowe (1984) presented a number of such *nonaccidental* properties of images that allow a strong inference to be made that that property is true of the three-dimensional region projecting the image. These are termed nonaccidental in that they are unlikely to be produced by an accident of a particular viewpoint. (Rock, 1983, has proposed a similar argument which he termed as "coincidence avoidance.") For example, if a section of contour is straight in the two-dimensional image, it is most likely produced by a straight contour in the three-dimensional world. The visual system appears to ignore the possibility that the straight edge in the image could be a projection of a curved edge from a viewpoint where it will project an image edge that is straight (Lowe, 1984). Other examples of nonaccidental properties are smooth curvature, symmetry, parallelism, and cotermination of edges. Biederman (1987) has derived a set of 36

Figure 2. An accepted grasp for chick sexing. (Modified from "Chick Sexing" by J. H. Lunn, 1948, *American Scientist, 36,* pp. 280–287. Copyright 1948 by the *American Scientist.* Photograph by the University of Minnesota Photographic Laboratory. Adapted by permission.)

primitive volumes that can be generated from contrasts of these nonaccidental properties. These may be sufficient to account for real-time object recognition (Biederman, 1987). Examples of properties of contour that do vary with viewpoint (and are hence *not* nonaccidental) are degree of curvature and length. Humans readily discriminate differences in nonaccidental properties (e.g., straight vs. curved) but are slow and inaccurate in judging metric properties such as length or degree of curvature. The general thesis pursued here is that, whenever possible, visual classification will be based on contrasts in nonaccidental properties (Biederman, 1987).

An Experimental Investigation of the Utility of the Contour Contrast

Was the simple contour difference providing the basis for a significant portion of the judgments of chicken sexers? We put this question to experimental test by assessing the effect of instructions that described the nature and location of the contour differences on the judgments of naive subjects. We did this by comparing the performance of these subjects before and after their instruction. We also compared the performance of the naive subjects with those of a small sample of professional sexers.

Method

Naive Group

A group of 36 students and faculty members from the University of California at Santa Cruz (UCSC) and the State University of New York (SUNY) at Buffalo, with no confessed experience or knowledge of chick sexing, viewed the 18 pictures in Figure 1. The pictures were arranged in random appearing order. Subjects were told that the pictures were divided equally between males and females and to use their "own intuitions to discriminate between the two types of genitalia." No accuracy feedback was provided. Two different arrangements of the 18 pictures were used with the arrangements balanced over subjects for the pre- and posttesting. Responses were recorded on a standardized answer sheet.

Following the pretest, half the subjects (those from UCSC) were given

an instructional sheet (Figure 3) that described the location and mapping of the critical contours. Approximately 1 min was required for this instruction. After reading the instructions, the subjects reclassified the 18 pictures, which were presented in a different randomized arrangement. The other half of the subjects (those from SUNY at Buffalo) retook the test without instructions or feedback from the first test. Subjects were told that the arrangements of pictures would not be the same for the pre- and posttest.

Professional Sexers

A total of five professional sexers were interviewed in their homes with a structured interview that assessed their training and employment history and use of perceptual and cognitive information in performing the task. Four of the five sexers had retired by the time of the interview. These four had been employed by Kimberly Farms in Fremont, California, which had been considered a leader in research and innovation in the poultry industry. Neither their interview responses nor their performance noticeably differed from the currently employed sexer. The professionals had been full-time sexers for a mean of 24.4 years (range = 18 to 36 years).

Pictures

The most important point to note about the pictures shown in Figure 1 is that they were initially created by Canfield (1940, 1941) to depict rare and difficult types. Their high level of difficulty provided us with an opportunity to obtain sufficient errors from the experts to correlate performance over the various pictures with the performance of subjects who received instructions in the course of our experiment. To our knowledge these are the only pictures of variations in chick genitalia. Not a single sexer, including one who went through a special school devoted to sexing, reported ever seeing pictures of chick genitalia.

Canfield (1941) reported the frequency of the various types in a sample of 10,000 chicks from "random commercial flocks." Male Types a and b were reported as constituting 64.32% of the sample; c, 7.85%; d and d', 0.58%; e and f, 20.53%; and h, 3.41%. For the females, a and b, constituted 57.22%; c, c' and d, 16.98%; e and f, 24.52%; and g and h, 1.27%. These frequencies were consistent with Mr. Carlson's spontaneous comments when he looked at Figure 1. For the males, d' was termed

Now you will be asked to sex-type another set of chickens. You should use the following rules to discriminate between the male and female genitals. The first part of your task should be to locate the informative region. In order to find this point, you should look for the two large cylindrical side lobes near the bottom of each picture. The genitals are located either between the ends of these two lobes (1) or slightly above the ends (2).
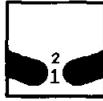
After you have found this area, you should study the swellings present there. It is at this point that males may be differentiated from females. Male chicken genitals tend to look round and fullish like a ball or watermelon. Here are two examples:

Female chicken genitals can take on two different appearances. They can look pointed, like an upside down pine tree, or flatish. Here are two examples:

Usually, but not always, male genitalia are larger. Sometimes either sex will appear to have double genitalia. You should differentiate the following 18 pictures based on this set of rules. Again, nine of the pictures are male and nine female in random order. You should begin the task by studying all of the pictures carefully. If you decide that a pictured set of chicken genitals belongs to a male, circle the "M" after the corresponding number on your answer sheet. If you decide that the pictured genitals belong to a female chicken then circle the "F" after the corresponding number on your answer sheet. Please work carefully. Once you understand these rules, you may turn this paper over and begin. Any questions?

Figure 3. The experimental instructions for chick sexing.

by Mr. Carlson as "rather rare," g as "rare," and h as "extra rare." For the females, a and e were termed "typical flat," and c′ as "difficult if (the bead) protrudes a lot." The frequencies allowed us to calculate expected performance levels.

Results

Interviews

Performance levels. The sexers estimated that an average of 2.4 months was required to reach a 95% accuracy level (range = 1.5 to 3.5 months). Their maximum performance averaged 99.4% (range = 99.3% to 99.5%) accuracy at a maximum rate that averaged 960 birds per hour (range = 900 to 1,000). Some of this time is spent getting new trays of birds and submitting the already sorted trays. Their estimates as to the amount of experience that they needed to reach their maximum performance levels ranged from 2 to 6 years (M = 3.3 years) with approximately one million birds sexed per year. Several of the sexers remarked that this was the time required to learn all the different "types." The variability in the estimates as to the degree of experience required to reach maximum performance was dependent on whether the sexer was referring to the maximum accuracy rate (generally achievable with less than a year's experience) or the time required to learn all the different (and very rare) types. The latter achievement required several additional years. These reports have considerable credibility. Careful records were kept as the sexers were paid by the number of birds accurately sexed.

Perceptual learning. In response to a question as to what was the most difficult part of the task, most sexers noted an aspect of image interpretation such as "reading the bead " or "different types." Only one sexer mentioned the manual manipulation as being the most difficult and this was phrased "so the bead could be read clearly."

Not a single sexer recalled being shown simple diagrams such as those in Figure 3 as part of their training. Training was accomplished by sexing live birds, which were then checked by experts. As we noted earlier, much of the training was based on learning a large number of types. Estimates as to the number of types varied; one sexer reported 45 of each sex, another reported that "they say there are over a 1,000 types for each sex."[1] Part of the motivation for mastery over the various types was that high accuracy rates were required (99% by some hatcheries) necessitating the accurate classification of rare configurations. In one hatchery in the 1960s, the individual sexers received 1¢ for each correctly sexed bird and were charged 35¢ for each error. Although the sexers said that they were looking for the bead, they also said that they were matching "types."

The sexers were keenly aware of their performance levels and highly desirous of improvement. An ambiguous bird would be examined by other sexers. Disagreements were resolved by *posting* the bird (cutting it open) on the spot and examining its sex organs. An alternative strategy was to bury possible mistakes by classifying ambiguous chicks as male.

Cognitive training. Every sexer interviewed noted the necessity of avoiding the "gambler's fallacy" (although none used that expression). When confronted with a sizable departure from an even split between males and females in a given tray, the sexers all reported resisting the impulse to classify the remaining difficult cases as members of the less frequent sex. Instead, if the sexers doubted their classification, they would reclassify *the entire batch.* All reported that they almost never altered their original classification. Although all reported that they could listen to music while working, four of the sexers reported that they could not "think of anything else" or talk while sexing a chick.

Experiment Results

The naive subjects averaged 60.5% correct (10.9 correct choices out of 18 pictures) in pretest and 84% correct in posttest, after instruction. The professional sexers averaged 72% correct choices. The improvement for the naive subjects was not a function of merely having a second opportunity to view the pictures. Performance on the second trial for the 16 naive subjects who did not receive instruction was 54.1%, a *decline* of 4.9% from their 59.0% accuracy on the first trial.

When performance was weighted in terms of the frequency of the pictures, the expected performance levels for the experts averaged 84.1% accuracy, the instructed subjects 89.8% accuracy, and the naive subjects, 65.3%. The experts never misclassified the most common types (pictures a and b for both sexes). All five professionals missed female c′, the one that Mr. Carlson termed "difficult."

[1] Indeed, the purpose of the Canfield (1940, 1941) articles was to illustrate the various types.

A mean accuracy score was computed for each picture for the three sets of data (those for the professional subjects and the pre- and posttests for the naive subjects). The Pearson product-moment correlation between the pretest naive subjects and the professionals was .21. The r between the posttest scores for the instructed subjects and the professionals was .82. This value was greater than the r of .63 for the pre- and posttest scores for the instructed subjects and even for a .57 correlation across a random split of the professionals. That is, after instruction the performance of the naive subjects more closely resembled that of the professional sexers than their own uninstructed performance a minute earlier. In contrast, Trial 1 to Trial 2 performance for the 16 naive subjects who received no instructional training was .87.

Discussion

More important than the improvement in performance of the naive subjects was that their item analysis more closely resembled that of the professional sexers as a consequence of the instruction. Before instruction, 4% of the variance of the professional sexers could be predicted by the naive subjects; after training this value increased to 67%. That is, after training the pictures that the nonprofessional group missed tended to be the same ones missed by the professional sexers.

We have thus demonstrated that a considerable portion of the visual learning in the classification of these pictures by expert sexers could be achieved by a brief instruction that merely described where a nonaccidental contrast in shape (concavity-convexity) could serve as the basis for classification.[2]

The 72% mean accuracy per picture in this study for the professionals sexers was markedly lower than the high 90% typically reported for on-the-job accuracy levels. About half of this difference could be accounted for by the higher miss rates for the rare and difficult types. The weighting for the frequency of type brought performance levels of the professional sexers up to 84%. It is our guess that the remaining discrepancy resulted from an inability to obtain the critical contour information from pictures. When confronted with an ambiguous live chick the sexer can introduce dynamic variation by alternately contracting and expanding the eminence by pressure variations. It is likely that such pressure makes the convexity–concavity (or flatness) distinction more readily discernible. Consequently, it was perhaps not surprising that most of the professional sexers voiced reservations about having to perform this task from pictures.

Another problem with the use of pictures was that in some cases, the reproduction and photography rendered a concave region convex in appearance. This happened in c' in Figure 1 in which the bead is aligned with the viewpoint so that it can be interpreted as convex. As noted earlier, all five sexers incorrectly classified that figure as a male. In some other cases, misleading variations in shading were introduced by the reproduction, and these might have lead to some erroneous judgments. However, the absence of color in the stimuli probably had no effect. Although some sexers and published articles mention that color can sometimes be used as a cue, none of the sexers reported that they used it regularly nor was it noted in the publications describing the Japanese methodology.

If much of the sexers' performance could be modeled as the
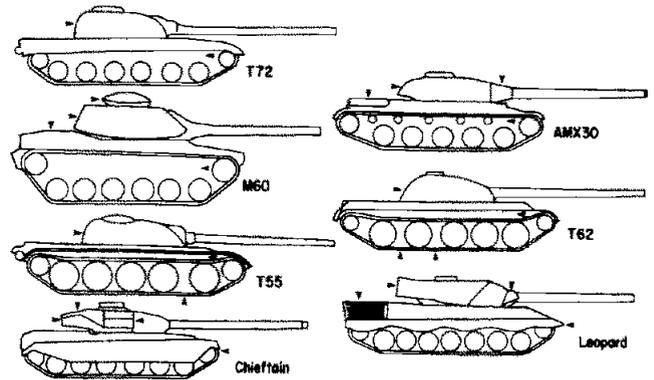

Figure 4. NATO and Warsaw Pact tanks. (Redrawn from *Comparison of Potential Critical Feature Sets for Simulator-Based Target Identification Training* by B. L. Kotas and D. W. Bessemer, 1980, Final Report, U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Knox Field Unit. The Warsaw Pact tanks, designated by the *T* in their names, all have completely rounded rear sections of turrets.)

use of a simple contrast in contour, why was there so much emphasis on learning types? It is likely that instance matching is needed to handle the (relatively rare) cases where the simple diagnostic contours of the eminence are not apparent. With these instances, accurate and relatively rapid classification may be based on contours that characterize that particular type but are not used for the most common instances. Such cases are instances of contingent information processing (e.g., Biederman, 1972) in that the relevant attributes for a particular stimulus depend on the status of other attributes.

The impressive gains in mimicking the performance of experts resulting from these minimal instructions suggests that learning might have been aided with pictures in which the critical features of the various types were indicated. The interviews revealed that a considerable portion of the learning over the 2.5- to 6-year period *following* the attainment of a high accuracy level was concerned with the learning of extremely rare types. It seems plausible that an album of such instances could lead to much more rapid mastery of these configurations by overcoming the problems of rarity and allowing immediate comparisons with similar types. This disproportionate allocation of effort to rare events is characteristic of a number of skills that require extremely high accuracy levels (e.g., diagnostic medicine or piloting planes).

The 62% correct choices of the naive subjects were above chance even when uninstructed. Most likely the reason for this level of accuracy was that the subjects' not-so-naive hypothesis about what might be diagnostic was partially accurate. For the most part, the presence of a prominent bead—which is the central singular structure—was interpreted as being male. This re-

---

[2] Note that our training was concerned only with the visual aspects of chick sexing. As in many such activities, chick sexing requires mastery of a number of subskills, including highly dextrous manual manipulation, which in turn requires sensitive tactile sensing. Consequently, despite equivalent performance on this task, it is highly unlikely that our trainees would be able to achieve immediate on-the-job performance levels that matched the professionals.

sulted in relatively high error rates when the bead was not large as with pictures c and g for males (in Figure 1) and when the bead was large (but concave) as in b and c' among the females. The above-chance interpretation of the bead among uninstructed subjects is a result similar to the finding that subjects who do not know Chinese can make above-chance discriminations of the meaning of a Chinese logograph (Koriat & Levy, 1979).

An Example From Another Domain: Classifying Tanks

One conclusion from this examination of chick sexing has been on the considerable advantage achieveable from a simple set of pictorial instructions that specify the location of diagnostic binary or trilevel contrasts of contour. As another example, consider Figure 4, which shows profile views of seven tanks, four from NATO countries and three from Warsaw Pact countries, adapted from training materials developed by Kotas and Besmer (1980). No doubt we have all seen TV footage of NATO maneuvers and May Day Parades, but few of us have spontaneously developed a perceptual model differentiating the two classes of tanks. NATO infantry personnel are required to determine friend from foe, and training is required. As suggested by the arrows of the Figure 4, the training largely consists of pointing out small nonaccidental differences (viz., characteristics that will be perceptible from a general viewpoint) in the contours between the two classes. A rule can be the following: If the rear of the turrent is completely curved (without notches) then it is a Warsaw Pact tank. From this one-sentence instructional session the reader can now perfectly categorize the two classes of tanks.

Note that these instructions were not of the metric details of degree of curvature or length of various features. More generally, the difficulty in spontaneously forming a class discrimination is that, with complex objects, the student does not know where to look or what to look for. Whenever possible, competent instructional materials for identifying members within the same basic level category, as with bird-guide books, explicitly tell the reader the location and nature of a distinguishing detail. Often, as with the case of birds, the details specify the color arrangements (or other surface features). But even here, the guides always note any available nonaccidental contrastive differences in contour in preference to metric information.[3]

The potential benefit for recognition speed in learning the characteristics of specific instances can be discussed with the case of tanks. It is likely that with additional training an observer will be able to determine not only the friend-or-foe status of a given tank but also the particular model. So when an experienced observer of tanks glimpses a Leopard, it is possible that the detection of the grill at the rear of the base allows faster classification of that tank as a Leopard (and thus a friend to NATO soldiers) than could be achieved by the classification based on the straight contour at the rear of its turret. With chickens, it is possible that the unique characteristics of specific types can be determined more rapidly and reliably than the contours of the bead that are diagnostic for the large proportion of the cases. We conjecture that such instance learning can be effective as long as the instances are not defined as conjunctions of independent attributes (Garner, 1974; Treisman & Gelade, 1980). For example, a conjunctive rule that could serve to distinguish the two U.S. tanks (the M60 and AMX30) from the others in Figure 4 could be *"six wheels on the ground* and *a curved hatch cover"* (the M60) or *"five wheels on the ground* and *a straight hatch cover"* (the AMX30). Such conjunctive rules are extraordinarily difficult to learn (Garner, 1974).

A contrast in a nonaccidental property can be readily learned and used as the criterion for rapid and accurate classification of complex objects. It is, of course, possible that such contrasts might not be available, in which case classification would have to be accomplished by prototype (or multiple-cue) matching. We suspect that nonaccidental contrasts will be spontaneously used whenever they are obvious. When not obvious because of small size, variability, or embedding in a complex object such as a chick cloaca or tank, a good instructional program is well advised to specify the contrasts rather than hope for their discovery.

---

[3] Greene, Pisoni, and Carrell (1984) observed that subjects learning to read speech spectrograms would use simple visual contrasts. Like the contrasts discussed in object perception, these appear to be nonaccidental (viz., invarient under slight changes in viewpoint) rather than metric. For example, the presence of a gap would be taken as an indicant for a consonant closure; a fuzzy patch would be interpreted as a fricative. It would be interesting to see if their 20-hr training period could be substantially reduced by the kind of instruction used in the present investigation.

References

Biederman, I. (1972). Human performance in contingent information processing tasks. *Journal of Experimental Psychology, 93*, 219–238.
Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115–147.
Canfield, T. H. (1940). Sex determination of day-old chicks. *Poultry Science, 17*, 235–238.
Canfield, T. H. (1941). Sex determination of day-old chicks II: Type variations, *Poultry Science, 20*, 327–328.
Carroll, J. (1984, October 22). A lifetime spent looking at chicks. *San Francisco Chronicle*, p. 4.
Garner, W. R. (1974). The processing of information and structure. New York: Wiley.
Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.
Greene, B. G., Pisoni, D. B., & Carrell, T. D. (1984). Recognition of speech spectrograms. *Journal of the Accoustical Society of America, 76*, 32–43.
Koriat, A., & Levy, I. (1979). Figural symbolism in Chinese ideographs. *Journal of Psycholinguistic Research, 8*, 353–365.
Kotas, B. L., & Bessemer, D. W. (1980). *Comparison of potential critical feature sets for simulator-based target identification training.* Final Report, U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Knox Field Unit.
Lowe, D. (1984). *Perceptual organization and visual recognition.* Unpublished doctoral dissertation, Department of Computer Science, Stanford University.
Lunn, J. H. (1948). Chick sexing. *American Scientist, 36*, 280–287.
Rock, I. (1983). *The logic of perception.* Cambridge: MIT Press.
Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.