International Journal of Comparative Psychology, Vol. 5, No. 4, 1992

INTERACTIVE MODELS OF COGNITIVE ABILITIES OF MONKEYS AND HUMANS (*Saimiri sciureus sciureus*; *S. boliviensus boliviensus*; *Homo sapiens sapiens*)

Roger K. Thomas
The University of Georgia

ABSTRACT: First, squirrel monkey and human data were complementary in validating the hypothesized difficulty of oddity and sameness-difference concept hierarchies. Second, both were used to refute the hypothesis that numerousness discriminations (e.g., 7 versus 8 items) require counting and to support the hypothesis that such judgments involve a prototype matching process.

This article summarizes two areas of investigation of cognitive abilities where comparable tasks have been used with humans and squirrel monkeys. The first area of investigation involves hierarchies of oddity and sameness-difference tasks that were developed to increase the precision of measurement within level 6 of an eight-level hierarchy (Noble & Thomas, 1985; Steirn & Thomas, 1990; Thomas & Frost, 1983). The second area of investigation involves the study of conceptual numerousness judgments, where the theoretical question of interest is to elucidate the mechanism by which such judgments are made (Terrell & Thomas, 1990; Thomas, Fowlkes, & Vickery, 1980; Thomas & Lorden, 1993; Thomas, Phillips, & Young, 1990). The question is whether counting is necessary or likely to be involved, or whether a simpler mechanism can be used to explain the judgments.

A Learning/Intelligence Hierarchy

The cognitive tasks can be considered in the context of an eight-level hierarchy of learning processes that are believed to by synonymous with

Address correspondence to Roger K. Thomas, Department of Psychology, The University of Georgia, Athens, GA 30602, USA.

the fundamental intellectual processes (Thomas, 1980). Briefly, the learning/intelligence hierarchy was constructed from Gagné's (1972) hierarchy of learning processes and from Bourne's (1970) approach to the study of concept learning in humans. The eight levels are (1) habituation and sensitization, (2) signal learning also known as classical or Pavlovian conditioning, (3) simple operant learning, (4) chaining units of simple operant learning, (5) concurrent discrimination learning, (6) learning *class concepts* which are based only on the logical operations, affirmation and negation, (7) learning *relational concepts* that involve using class concepts in conjunctive, disjunctive, or conditional relationships or their respective complementary operations, and (8) learning relational concepts that involve using class concepts in biconditional relationships or its complementary operation, the exclusive disjunction.

Thomas's (1980) construction of the hierarchy included the following modifications of Gagné's (1972) and Bourne's (1970) approaches: (a) habituation and sensitization were added below signal learning, the bottom of Gangé's hierarchy, and (b) the concept learning hierarchy based on Bourne (1970; levels 6–8 above) was substituted for Gagné's highest three levels. Gagné's levels 6–8 were Concept Learning, Rule Learning, and Problem Solving, and they were defined by tasks that in most cases are likely to be unique to human abilities owing to their verbal or mathematical content. Thomas's substitution is justifiable because the concept learning hierarchy (levels 6–8 above) is fundamental to Gagné's levels 6–8; that is, the tasks that helped define Gagné's levels 6–8 can be reduced to the logical operations in levels 6–8 here.

Thomas (1980) also added an operational distinction between *absolute* and *relative* class concepts at level 6. Namely, the defining attributes of absolute class concepts are inherent in each discriminandum (e.g., each tree manifests its "treeness"). In contrast, the defining attributes of relative class concepts are not inherent in the discriminandum manifesting the correct choice but represent a relative property among the discriminanda (e.g., "oddity" as manifested in a triangle versus two circles or as manifested in a circle versus two triangles). Further details may be seen in Thomas (1980) and Steirn and Thomas (1990).

The hierarchy offers powerful advantages for comparative studies of intelligence. To summarize: (a) theoretically, the hierarchy can be used with any species from protozoans to humans, (b) it can be used relatively independently of the confounding effects of contextual variables such as sensory and motor differences, differences in motivation, etc., because it assesses processes and does not depend on specific tasks; that is, each assessment task can be adapted to each species' unique sensory, motor, motivational, etc., requirements, (c) it encompasses all of the fundamental learning processes, (d) it can be applied retroactively or proactively to any learning study whether or not the study was planned in the context of the hierarchy, and (e) it can be used for ontogenetic or phylogenetic comparisons.

It is likely that all vertebrates are capable of at least some degree of success in concurrent discrimination learning, level 5 (see Table 4, Thomas, 1986). It is also likely that cognitive and intellectual differences among mammals will be found among the *conceptual* abilities represented at levels 6–8. Owing to the latter, the research in my laboratory has been focused on the abilities of nonhuman animals to learn to use class and relational concepts.

Oddity and Sameness-Difference Concepts

There are systematic ways to increase the cognitive demands and, therefore, the precision of measurement within levels 7 and 8, and sublevels of 7 and 8 can be developed without theoretical limit (see Tables III and IV, Thomas, 1980). However, ways of distinguishing abilities within or between species within level 6 are not immediately obvious. Since it seems likely that many species will reach a ceiling within level 6 (i.e., be able to use class but not relational concepts) it seems essential to find systematic ways to increase the precision of measurement within level 6. We have investigated two related types of tasks, oddity and sameness-difference (hereafter, SD).

With one exception that is relatively minor and will be discussed below, systematic hierarchies of oddity and SD tasks can be constructed by manipulating discriminative cues based on attributes such as color, form, and size. The cues can be manipulated to be (a) *relevant* to, (b) *constant* and, therefore, uninformative, or (c) *ambiguous* to the discrimination being required. By ambiguous, it is meant that the cues vary in an uninformative way. While being uninformative, constant cues are not distracting, but ambiguous cues are likely to distract.

The aforementioned exception to the systematic construction of oddity or SD hierarchies occurs when one makes the transition from *identical* nonodd or sameness discriminanda to nonodd or sameness discriminanda that are *not identical* but have more attributes or properties in common than they have with the odd or difference discriminanda. In this instance, but only at the transition point, the increasing difficulty is not determined logically but must be determined empirically. This will be discussed further below.

For simplicity, the following discussion is only in terms of oddity problems. However, everything that is written here about oddity problems, except as noted, can be applied equally to SD problems provided "odd" and "difference" are used interchangeably and provided "nonodd" and "sameness" are used interchangeably.

In a typical three-discriminanda oddity task, each trial consists of presenting the odd discriminandum and the two nonodd discriminanda.

For example, one might use two red cubes and one blue ball where the cubes are the same size but are smaller than the ball. In this example, the color, form, and size cues are all relevant; that is, each can be used to distinguish the odd from the nonodd discriminanda. Alternatively, one might hold one or two of the cues constant, thereby leaving two or one relevant cues, respectively. While, empirically, the difference between the odd and nonodd discriminanda based on one cue alone might be as obvious and easy to discriminate as when all three cues are relevant (as seems to be the case for squirrel monkeys and humans), theoretically, a problem with three relevant cues should be easier than a problem with two or one relevant cues. Of course, it is possible and seems likely that some species or younger subjects within a species will find that a one-relevant-cue problem is more difficult than a two- or three-relevant-cue problem.

When the transition is made such that the discriminanda no longer include identical nonodd objects, the discriminations that are required become more difficult. For example, two cubes might be the same smaller size compared to one larger ball (thus, size and form remain relevant) but the objects might each be of a different color. In our research, we have varied randomly from trial to trial which types of cues will be relevant, constant, or ambiguous. To conclude with these examples, we might have a small red cube, a medium-sized blue pyramid, and a large red ball on one trial (thus, color determines the odd object) and then have a small green pyramid, a middle-size red ball, and a large blue ball on the next trial (thus, form determines the odd object). It is obvious that such discriminations are of a different order of difficulty compared to those where the nonodd objects are identical.

Theoretically, there should be increasing cognitive demands and, therefore, increasing performance difficulty as one goes from three cues to two to one relevant cue or as one goes from one to two ambiguous cues, etc., even if empirically such changes do not always challenge all species. As suggested earlier, the logical structure of the hierarchy falters when one changes from identical to nonidentical nonodd or sameness discriminanda. The logical progression from levels 1 to 3 (based on decreasing relevant cues) and from levels 4 to 6 (based on decreasing relevant cues and increasing ambiguous cues) seems clear. However, logically it is not possible to say whether problems with one relevant and two constant cues would be easier or harder than problems with two relevant and one ambiguous cues. Thomas and Frost (1983) hypothesized that the ambiguous cue would cause more difficulty and designated the two-relevant, one-ambiguous cues problems as being at level 4 leaving the one-relevant and two-constant cues problems to be at level 3. Therefore, among our goals was to determine empirically whether the oddity hierarchy as constructed would lead to systematic performance differences.

Before discussing our empirical research, it is useful to add the following two general points. First, up to now the oddity (and SD) hierarchies have been discussed in terms of varying color, form, and size. By varying these three properties, one can construct a 6-level hierarchy (see illustrations in Steirn & Thomas, 1990). By adding another property such as number, one can construct a 10-level hierarchy (an example of a trial might be: two small red balls are one discriminandum, two small red balls are another, and three large blue cubes is the third discriminandum on one particular oddity trial). Other properties could be added (e.g., placing the pairs of small red balls on a white square and placing the three large blue cubes on a white circle) making a 15-level hierarchy. Additional properties could also be introduced and hierarchies with more levels could be constructed.

Second, while the oddity of SD hierarchies can be constructed in a highly similar manner, there is a significant conceptual difference between them. Namely, as long as a pair of discriminanda manifesting sameness consist of identical objects, the discrimination between the sameness and difference pairs of objects can be made as an absolute class concept; that is, it is not necessary for an animal to compare the sameness and difference discriminanda in order to affirm which pair of objects manifest sameness and which manifest difference. It is only when the SD tasks involve nonidentical sameness pairs that comparison between discriminanda becomes necessary. Oddity, however, always requires comparison and is, therefore, always a relative class concept.

The difference between absolute and relative class concepts is important, because nonprimate animals have succeeded on tasks requiring the use of absolute class concepts, but owing to methodological confounds it is unclear whether any nonprimate animal has ever succeeded on a task that requires the use of a relative class concept (see Steirn & Thomas, 1990, for further related discussion). If the capability of learning absolute and relative class concepts represents a "breakpoint" in phylogenetic cognitive development, then having a series of tasks (viz., the SD hierarchy) which makes the transition from absolute to relative class concepts in a systematic way is exceptionally valuable.

Empirical Tests of the Oddity and SD Hierarchies using Squirrel Monkeys and Humans. Thomas and Frost (1983) trained squirrel monkeys on oddity tasks beginning with the level 1 task and proceeding in succession to the level 6 task. Training on the level 1 task was not limited in the number of trials, but all monkeys met the joint-criterion (see below) in 1,200 trials or less. Since succeeding tasks should involve considerable transfer of training from level 1, a limit of 400 trials (ten sessions of 40 trials each) was planned if they failed to meet the joint criterion of 36 correct in a 40-trials session and a significant run of errorless trials ($p < .01$).

Details of the monkeys' performances may be seen in Thomas and

Frost (1983; especially Table 2). In summary, we found that, as predicted, each succeeding level was more difficult than the previous one, except that, contrary to prediction, level 4 (the first with ambiguous cues) was easier than level 3 and was as easy as levels 1 and 2.

However, we had overlooked a reasonable noncognitive explanation. Male squirrel monkeys are deficient in color vision (Jacobs & Neitz, 1985), and level 3, as noted above, had one relevant cue while level 4, even with its ambiguous cue, had two relevant cues. Given that by chance the relevant cue at level 3 would be a color cue on one-third of the trials, the monkeys' color vision deficiencies likely accounted for their poorer performances on level 3. This also explains in part their poorer performances on levels 5 and 6 which, like level 3, had only one relevant cue. However, color vision deficiencies do not provide the complete explanation, because their performances on level 6 were worse than those on level 5.

In view of the confounding role of color vision deficiencies in the squirrel money, Noble and Thomas (1985) decided to test the empirical validity of the oddity hierarchy using humans. After screening the human participants for color vision deficiencies, Noble and Thomas (1985) tested each person on only one of the oddity tasks; 10 people were tested on each task. Generally, the tasks at all six levels were too easy for our adult humans to provide clear differences in performance at each level. Nevertheless, the data provided support for the hierarchy. For example, level 6 required more mean trials to criterion than level 5, although the difference was not significant, but both levels 5 and 6 differed significantly from each of levels 1–4, etc.

Most importantly, in view of the unexpected but explainable (in terms of color vision) finding that the monkeys found level 4 easier than level 3, there was evidence that the humans found level 4 more difficult than level 3. Specifically, their response latencies were significantly longer at level 4 than at level 3, and they approached statistical significance ($p < .08$) in taking more trials to criterion on level 4 than on level 3.

Our most recent attempt to validate both the oddity and SD hierarchies using humans involved manipulations that were intended to increase task difficulty in a way that might provide clearer validation of the predicted differences between successive levels (Steirn & Thomas, 1990). Specifically, each person was trained on a random sequence of trials that were administered concurrently from *three* levels (either 1–3 or 4–6). This was done for some using oddity problems and for others using SD problems.

Despite the added complexity of mixing trials from three task levels, Steirn and Thomas (1990) reported the same general findings on the oddity tasks that were reported by Noble and Thomas (1985). For example, differences in percentages correct and in response latencies as a function of task level were usually in the predicted direction, but the differences were not usually significantly different between successive levels. Comparable to Noble and Thomas's findings on the oddity tasks, response latencies were significantly longer on level 4 trials than on level 3 trials. Although the difference in percentages correct between levels 3 and 4 were in the predicted direction, whey were not statistically significant.

Important findings emerged on Steirn and Thomas's (1990) SD tasks. The differences between levels 3 and 4 on both the percentage correct measure and the response latency measure were statistically significant in the predicted direction. This is important because, as discussed earlier, levels 1–3 in SD tasks can be done on an absolute class conceptual basis but levels 4–6 require that they be done on a relative class conceptual basis.

Obviously, in view of the tasks being generally too easy for adult humans and in view of the male squirrel monkey's visual problems, further research on the oddity and SD hierarchies is needed. Future research might be done with young humans or with primates that have trichromatic color vision that could validate and expand the usefulness of the oddity and SD hierarchies.

Conceptual Numerousness Judgments

The typical numerousness judgment task used with animals has been to display arrays of entities (e.g., black dots on a white background) and have the animal discriminate between an array with one number of entities and an array with another number of entities. There has long been an interest in animals' abilities to make such judgments (Honigman, 1942; Salman, 1943; Wesley, 1961), but Wesley concluded that only Hick's study (1956) had been sufficiently free of confounding variables to conclude that the animals' judgments were based on numerousness. Hicks reinforced rhesus monkeys (*Macaca mulatta*) for choosing arrays of 3 items versus arrays of 1, 2, 4, or 5 items. While there have been some well controlled studies (Davis & Pérusse, 1988; Thomas & Lorden, 1993) poorly controlled studies have been prevalent.

Typical confounding variables in the early studies were (a) having dots of uniform size where cumulative area or differential brightness were possible discriminative cues and (b) failing to control for the odor of the reinforcers which may have cued the animal to the correct choice. Thomas and Lorden (see Table 1; 1993) listed these and other possible confounding cues or interpretations that must be avoided before attributing numerousness judgments to animals.

Thomas, Fowlkes, and Vickery (1980) incorporated the appropriate controls and used systematic training procedures to determine the squirrel monkey's likely upper limit in ability to discriminate consecutive numerousness arrays. Both monkeys discriminated seven versus eight dots at a high level of success (90% based on 45 correct in a 50-trials session; hereafter, the form 7:8 will be used to describe such discriminations). One of the monkeys met the 90% criterion on 8:9 but failed to meet criterion on 9:10 within the preset limit of 500 trials, although he performed at a level of about 75% correct on the 9:10 task.

Subsequently, Terrell and Thomas (1990) used the number of sides of randomly constructed polygons as discriminanda. Their monkeys' best performances were that two of four monkeys met a 90% correct criterion on 7:8 (36 correct in a 40-trials session), one met the criterion on 6:7, and one met the criterion on 5:7.

Processes to Explain Numerousness Judgments. Gallistel (1988, 1990) has argued forcefully that numerousness judgments of dot arrays are based on counting. He has strongly opposed processes such as the prototype matching one that we propose to explain such judgments. Central of Gallistel's argument is the reported serial increase in response times as the number of entities in an array increases. According to Gallistel (1988), citing data from Mandler and Sheebo (1982):

> It takes 30 msecs longer to recognize twoness than to recognize oneness, 80 msecs longer to recognize threeness than twoness, 200 msecs longer to recognize fourness than threeness and from fourness on up there is an increment of 350 msecs per item (p. 586).

However, not all studies have shown serial increases in response times, and the procedures used by Mandler and Sheebo (1982) may not be applicable to explain the results of Thomas et al. (1980) and Terrell and Thomas (1990).

An alternative explanation to counting that Terrell and Thomas (1990) proposed was that their monkeys acquired prototypes for absolute class concepts such as "threeness," "sevenness," etc. and then applied them to make accurate numerousness judgments of new arrays of entities. That 7:8 emerged as a common upper limit in both the dot and sides-of-polygons studies suggested the possibility of an underlying common process.

The upper limit of numerousness prototype acquisition and use is probably related to information processing channel capacity as exemplified in Miller's (1956) well known "magical number seven, plus or minus two." Individuals can then learn to use the acquired prototypes to affirm the numerousness of *new* arrays of entities by matching each new array with its numerical prototype. Monkeys will likely require more trials to acquire and use prototypes than humans, because monkeys lack the prior experience with number that is typical of humans. Based on our findings with monkeys and humans there is no reason to think that the monkeys' channel capacity for numerousness will differ significantly from that of humans. After discussing why we proposed a noncounting process explanation, I will report our findings from an experiment that we conducted to test the response latency hypothesis.

Terrell and Thomas (1990) proposed a noncounting explanation of our monkey data, because our monkeys did not have the requisite experience and skills to count. We based the requisite skills on three of the five principles of counting presented by Gelman and Gallistel (1978), namely, the one-to-one correspondence, stable-order, and cardinal principles. Gallistel later (1990) indicated that the fourth and fifth principles, order-irrelevance and abstraction, are not essential to demonstrate counting, a position I have supported (Thomas, 1992).

One-to-one correspondence requires that the individual apply unique tags (such as but not limited to Arabic numerals) to the entities being counted. Gelman and Gallistel (1978) apparently did not require direct evidence that the individual had acquired or used the tags nor, therefore, that they had applied them in stable order. However, direct evidence is necessary, because there is an alternative noncounting process, prototype matching, that can explain the animals' numerousness judgments. Our monkeys had no opportunity to learn or apply tags, and most importantly, they did not need to learn and use tags, because the prototype matching process is simpler than counting. Prototype matching is also consistent with the way other kinds of class concepts are likely to be acquired and used by animals.

Empirical Tests of Counting Versus Prototype Matching Explanations. We (Terrell & Thomas, 1990; Thomas & Lorden, 1993; Thomas, Phillips, & Young, 1990) hypothesized that if an individual can make a discrimination between two arrays accurately (as defined by being correct on 90% of a set of trials), the judgments for the two arrays will be made with similar response latencies. Additionally, according to Gallistel and based on the data in the quotation above, it would require (x + 1,360) msecs to judge sevenness and (x + 1,710) msecs to judge eightness. The value, x, represents the time for the oneness judgment that was not specified in the quotation from Gallistel.

Since the data needed to examine the relationship of response times to numerousness judgments could be obtained more easily with humans than with squirrel monkeys and since humans were the subjects in the studies on which Gallistel (1988, 1990) based his arguments, we (Thomas et al., 1990) elected to test the response latency hypothesis by using humans. We used both dots and sides of polygons as the numerousness discriminanda. Since the findings were generally consistent for both dots and sides of polygons and in the interests of brevity and clarity, most of the discussion here will be limited to the dots.

Each subject was tested first on the 3:4 number-pair and then on another number-pair discrimination, either 4:5, 5:6, 6:7, 7:8, 8:9, 9:10, or 10:11. If dots were tested first, then the person was subsequently tested on polygons and vice versa. Twenty people were tested on each number-pair (except, of course, all 140 subjects were tested on 3:4).

A numerousness array was presented tachistoscopically for 200 msec and was followed immediately by a masking stimulus the use of which was intended to prevent counting based on afterimages. The dots were varied in size so cumulative area and differential brightness could not be used reliably as discriminative cues, and all trials consisted of unique arrays so pattern learning was precluded.

The person was told before each set of trials which of two numbers would be presented. Single exemplars of a number-pair were presented one at a time in a randomized order, and the individual was instructed to press a button that corresponded to its number as quickly as possible. Participants were trained to a criterion of 90% correct in a block of 20 consecutive trials or until a maximum of 200 trials per number-pair had been administered. The following results are based only on those who reached the 90% criterion.

Our specific prediction was that the response times for *consecutive* arrays would not differ, but there were no significant differences in mean response times for *any* array of numbers from 3 to 11. Mean response times ranged from 720 msecs for arrays of 10 (in the 9:10 problem; based on the 3 of 20 people who reached criterion) to 1,010 msecs for arrays of 11 (in the 10:11 problem; based on 2 of 20 participants). As might be expected there was a general decrease in the number of people who reached criterion as the size of the number-pair increased. For example, all 140 met the criterion on 3:4 and 20 met the criterion on 4:5, but only the few noted above met the criterion on the 9:10 and 10:11 problems. Also, as might be expected, there was a general increase in the mean number of trials to criterion as the size of the number-pairs increased from 30 trials on 3:4 to 170 trials on 10:11.

According to Gallistel (1988, 1990), we should have found response times that increased serially as the number of entities in an array increased, but we did not find significant differences in response times across our arrays. Gallistel also predicted response times that were considerably longer than we obtained. According to Gallistel, we should have found response times of approximately (x + 2,760) msecs for arrays of 11 dots but our mean response time for 11 dots was only 1,010 msecs. Since the serial increase in response times was central to his justification for the counting explanation, our data refute Gallistel's basis for the counting explanation.

As a closing point in support of the prototype matching process, I urge the reader to experiment informally with prototype acquisition and use. For example, one can easily view at a glance (or view with the deliberate intention to avoid counting) arrays of items, such as clusters of farm animals in a field as one passes in an automobile or train, and try to guess accurately the number of items. Of course, one can then count the items to confirm the number. I believe that you will discover that your judgments for small arrays (e.g., 3, 4, and 5 which are probably learned passively during one's lifetime) are usually accurate and that with practice you will increase the number of items in arrays that can be judged accurately. Based on our findings, most of us will peak in accuracy at about 6:7 for consecutive numerousness judgments, but there are occasional savants who acquire and use prototypes for numbers as high as 10 and 11 or possibly higher. Practice may improve skill up to a limit, but I suspect that limit is within a small extension of Miller's (1956) magical number 7 plus or minus 2. Neither our research so far nor this informal experiment address discriminative numerousness judgments that are nonconsecutive (e.g., 25 vs. 50 items) and that can likely be learned accurately as well.

Conclusion

The present research has suggested the feasibility of developing and testing models of cognition and intelligence that can be applied to human and nonhuman animals. While the usual direction of animal-human model research has been to seek an animal model to test some question of consequence to humans, we have used humans as models to investigate issues that arose in animal research. Of course, the fundamental processes addressed in our research apply equally to humans and nonhuman animals.

ACKNOWLEDGEMENT

The author thanks Cheryl D. Young for her helpful suggestions on the manuscript.

REFERENCES

Bourne, L. E. (1970). Knowing and using concepts. *Psychological Review, 77,* 546–556.
Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences, 11,* 561–579.
Gagné, R. M. (1972). *The conditions of learning.* New York: Holt, Rinehart & Winston.
Gallistel, C. R. (1988). Counting versus subitizing versus the sense of number. *Behavioral and Brain Sciences, 11,* 585–586.
Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: The MIT Press.
Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number.* Cambridge, MA: Harvard University Press.
Hicks, L. H. (1956). An analysis of number-concept formation in the rhesus monkey. *Journal of Comparative and Physiological Psychology, 49,* 212–218.
Honigmann, H. (1942). The number conception in animal psychology. *Biological Review, 17,* 315–337.
Jacobs, G. H., & Neitz, J. (1985). Color vision in squirrel monkeys: Sex-related differences suggest the mode of inheritance. *Vision Research, 25(1),* 141–143.

Mandler, G., & Sheebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111,* 1–22.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review, 63,* 81–97.

Noble, L. M., & Thomas, R. K. (1985). Oddity and dimension-abstracted oddity in humans. *American Journal of Psychology, 98,* 549–557.

Salman, D. H. (1943). Note on the number conception in animal psychology. *British Journal of Psychology, 33,* 209–219.

Steirn, J. N., & Thomas, R. K. (1990). Comparative assessments of intelligence: Performances of *Homo sapiens sapiens* on hierarchies of oddity and sameness-difference tasks. *Journal of Comparative Psychology, 104,* 326–333.

Terrell, D. F., & Thomas, R. K. (1990). Number-related discrimination and summation by squirrel monkeys (*Saimiri sciureus sciureus* and *Saimiri boliviensus boliviensus*) on the basis of the number of sides of polygons. *Journal of Comparative Psychology, 104,* 238–247.

Thomas, R. K. (1980). Evolution of intelligence: An approach to its assessment. *Brain, Behavior, and Evolution, 17,* 454–472.

Thomas, R. K. (1986). Vertebrate intelligence: A review of the laboratory research. In R. J. Hoage & L. Goldman (Eds.), *Animal intelligence: Insights into the animal mind* (pp. 37–56). Washington, DC: Smithsonian Institution Press.

Thomas, R. K. (1992). Primates' conceptual use of number: Ecological perspectives and psychological processes. In T. Nishida, W. C. McGrew, P. Marler, M. Pickford, & F. B. M. de Waal (Eds.), *Topics in primatology: Volume 1. Human origins.* Tokyo: Tokyo University Press.

Thomas, R. K., & Frost, T. (1983). Oddity and dimension-abstracted oddity (DAO) in squirrel monkeys. *American Journal of Psychology, 96,* 52–64.

Thomas, R. K., & Lorden, R. B. (1993). Numerican competence in animals: A conservative view. In S. Boysen & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 127–147). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Thomas, R. K., Phillips, J. A., & Young, C. D. (1990, November). *Numerousness judgments by Homo sapiens in a species-comparative context.* Paper presented at the meeting of The Psychonomic Society, New Orleans, LA.

Thomas, R. K., Fowlkes, D., & Vickery, J. D. (1980). Conceptual numerousness judgments by squirrel monkeys. *American Journal of Psychology, 93,* 247–257.

Wesley, F. (1961). The number concept: A phylogenetic review. *Psychology Bulletin, 53,* 420–428.