Overcoming contextual variables, negative results, and Macphail's null hypothesis

Roger K. Thomas
Department of Psychology, University of Georgia, Athens, Ga. 30602

Macphail discounted my approach to assessing comparative intelligence as well as Barlow's (1983) and Bullock's (1986), noting that such schemes "cannot be accepted until experimenters consistently report failures by certain species to master specific tasks." This is an unfair condition owing to the well-known difficulty in publishing "failures" or negative results. Typically, the difficulty is due to the confounding effects of contextual variables (e.g., sensory, motor, and motivational ones) on the assessment of intelligence. Discounting my approach overlooks its strength in avoiding such contextual confounding. After summarizing my approach, which offers a way to reject Macphail's "null hypothesis," I will consider perhaps the most widely used "complex" task with vertebrates as an example involving the issue of publishing negative results in the context of efforts to reject his hypothesis.

My scheme (Thomas 1980) posits an eight-level *hierarchy* of types of learning; capacity to learn was equated with intelligence. *All* learning (laboratory and natural habitat) may be reducible to components of the hierarchy. The eight levels are (1) habituation and sensitization; (2) classical conditioning; (3) operant conditioning; (4) chaining; (5) concurrent discrimination learning; (6) responding to absolute and relative class concepts; (7) responding to conjunctive, disjunctive, and conditional concepts, and (8) responding to biconditional concepts. (See Thomas [1980] for rationale, definitions, and caveats.)

The learning-hierarchy approach avoids confounding contextual variables, because the sensory, motor, motivational, and other aspects of exemplary tasks at each level are adapted to each species. The same kinds of discriminanda, response demands, and incentives used at one level can be used at other levels. Then, for example, if an animal succeeds at one level but fails at another, the reason would be the intellectual demands of the task rather than contextual variables. An exception might be the onset of fatigue, but that could be determined by returning to a lower level task.

Meaningful sublevels are possible beginning with level 4. This will be illustrated with the oddity problem, a level 6 task, because it is relevant to comments below. The order of the alpha-numeric "names" shown below reflects an order of task difficulty (Noble & Thomas 1985; Thomas & Frost 1983). Typically, the oddity problem involves one odd and two identical, nonodd stimuli. With color, shape, and size varying, the odd stimulus might share no attributes with the nonodd stimuli which, when identical, share all three (the O0N3 task), or the odd stimulus might share one or two attributes with the nonodd stimuli (O1N3. O2N3). In "dimension-abstracted oddity," the nonodd stimuli are not identical but have more attributes in common than with the odd stimulus (e.g., O0N2, O1N2, O0N1).

The oddity problem has probably been used with more different species of vertebrates than any other "complex" problem. Except for studies using nonhuman primates and one study using pigeons, claims for the use of the oddity concept by other nonhuman animals do not withstand close scrutiny (Thomas & Frost 1983). The issue is whether it was possible that the animal learned associations between the *specific* odd stimuli and reward or rather responded to oddity conceptually. The best control for specific versus conceptual responding is to analyze only the first-trial performance on new oddity problems. Despite the aforementioned criticism and control having been noted several times at least since 1948, a number of recent studies using pigeons unjustifiably claim that the oddity concept has been used. Some investigators were surely aware of the criticism and control before they published. My guess is that lurking in some of their files are data showing failures by pigeons to respond reliably to the odd stimuli on the first trials.

The exception among the pigeon studies was Lombardi et al.'s (1984) which used oddity test problems with reinforcement always withheld; this was mixed in with other oddity problems (presented repeatedly) in which reinforcement was administered, and the pigeons chose the odd stimulus at better than chance frequencies on the test problems. However, pigeons are unlikely to succeed on the higher-level oddity problems described above, and the authors of such studies are unlikely to publish their negative results.

Similarly, the few studies assessing the rat's use of the oddity concept are inconclusive, including those which claimed positive results. My recent effort to publish a "negative" result was rejected for nonmethodological reasons ("Visual and olfactory oddity learning in rats: What evidence is necessary to show conceptual behavior?") One reviewer and the editor mentioned the "negative" results. The rats received a total of 300, five-trial oddity problems. The rats learned to respond better than chance on the second trials of new problems early in training but remained at chance on the first trials throughout. This showed that sensory, motor, and motivational variables did not account for the failures on trial one and suggested that success on trial two was due to learning quickly what specific stimulus and reward contingencies were in effect on each five-trial problem. In other words, they showed evidence for learning set formation but none for knowledge of the concept of oddity. I will persist in publishing this study, but how many might not?

None of this commentary diminishes Macphail's main point that, so far, the data do not refute his null hypothesis. My approach to the comparative assessment of intelligence offers a way to reject it – provided the "failures" necessary to reject it become part of the published record.