COL·LECCIÓ
HOMENATGES 51

# Associative Learning and Cognition

## Homage to Professor N. J. Mackintosh. «In memoriam» (1935-2015)

Edited by
J. B. Trobalon, V. D. Chamizo

UNIVERSITAT DE BARCELONA

# *Contents*

# *Acknowledgments*

The chapters published in this volume are a homage to Professor N. J. Mackintosh (1935-2015), an outstanding academic and a dear friend and colleague to all of the participants. The topics have been freely chosen by the authors. The fact that this book appears in a specific collection ("Homages") of the publishing section of the University of Barcelona deserves some explanation. Professor Mackintosh collaborated with different members of the Departament de Psicologia Bàsica (at present, Departament de Cognició, Desenvolupament i Psicologia de l'Educació), Universitat de Barcelona (UB), from the beginning of the 1980s until he passed away (on 8th February, 2015), after a brief illness. We were all devastated by the news. For many years our collaborative research aimed to see whether the spatial and the temporal domains share the same or similar conditions, basic effects, and mechanisms. Our results showed that many of the phenomena found in experiments on Pavlovian and instrumental conditioning and simple discrimination learning were also observed in our laboratory in experiments where rats were required to locate a goal by means of two or more distal landmarks. These phenomena included: blocking, overshadowing, latent inhibition, perceptual learning, and changes in attention to relevant and irrelevant cues (see the Appendix section of this volume for these references). Standard associative theories could explain all these phenomena (Rescorla & Wagner, 1972; Mackintosh, 1975; McLaren, Kaye, & Mackintosh, 1989; Pearce & Hall, 1980; Wagner, 1981). All the previous results are inconsistent with O'Keefe and Nadel's original proposal (1978) that locale learning (i.e., behaviour based on a representation of allocentric space, or cognitive map) occurs non-associatively, in an all-or-none manner, and that animals constantly update their cognitive map of their environment.

During the last years the main part of our collaborative research aimed to see whether male and female rats trained in a triangular shaped pool to find a hidden platform whose location was defined by two sources of information — one particular corner of the pool and a salient landmark positioned immediately above it — differed in their preferred mode of solution (geometry of

the pool cue vs. landmark cue) and also in the amount they learned about these two cues. Our results (see the Appendix section) are in agreement with previous findings showing that males and females do not always use the same strategies when solving spatial tasks (Williams, Barnett, & Meck 1990), age being a critical factor. In addition, males and females do not learn the same about the two sources of information. A biological origin?

Our collaboration with Professor Mackintosh began in 1982, on the occasion of a stay of V. D. C.* in the Department of Experimental Psychology (nowadays, Department of Psychology) as a postdoctoral student attending a course on "Animal Learning and Physiological Psychology", which was granted by the European Science Foundation (an ETP twinning grant). For many years Professor Mackintosh was a formal member of the UB research group *Learning and Cognition: A Comparative Approach* (www.gracec.info/). This collaboration produced a considerable number of contributions to meetings, publications, granted research projects, and research stays by different members of this group in the Department of Psychology at Cambridge University. During these years Professor Mackintosh was a frequent visitor to our university. It is worth mentioning his participation as an Invited Speaker in various courses organized and subsidized by different Catalan institutions (UB Department of Basic Psychology, UB Faculty of Psychology, financial help from the UB Chancellor, and UB Institute of Education Sciences — ICE abbreviation in Spanish), and in various PhD tribunals as an external examiner. However, his most important legacy to the UB was his outstanding contribution to the joint publications during those years that are included in the Annex section of this volume. (For some information about his distinguished career see Miller, 2016 and Hall, 2016 in this volume; visit also www.psychometrics.cam.ac.uk/about-us/directory/nick-mackintosh.)

As a sign of recognition and gratitude, Professor Mackintosh received posthumously (11 November, 2015) the highest honor, a Gold Medal, that the University of Barcelona can give to a person who is no longer with us. His widow Leonora Brosan Mackintosh, Lee, collected it in his name from the Chancellor of the University (Professor Dídac Ramírez), in a moving ceremony held in the Main Hall (the "Paraninfo") of the Historic Building (www.ub.edu/ubtv/video/acte-homenatge-prof-mackintosh). Several of his children (as well as other friends and colleagues) also attended the event, which was followed

* At the University of Barcelona since 1980.

by a formal lunch. At the ceremony, this book was announced to commemorate the first anniversary of the Gold Medal prize. In fact, when Professor Mackintosh died he was a candidate to be nominated Doctor Honoris Causa from the University of Barcelona. That process was automatically stopped because a requirement at our university is that the nominee personally attends a formal ceremony to collect the prize.

Thank you so much Nick!

The editors would like to express their deep gratitude to all the authors that have made this book possible. We should also like to express our gratitude to Lucy Mackintosh, for excellent language review in some chapters.

V. D. Chamizo, J. B. Trobalon
June 2016

## References

Mackintosh, N. J. (1975). A theory of attention. Variations in the associability of stimulus with reinforcement. *Psychological Review*, *82*, 276-298.

McLaren, I. P. L., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 102-130). Oxford: Oxford University Press.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the efectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale: Erlbaum.

Williams, C. L., Barnett, A. M., & Meck, W. H. (1990). Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. *Behavioral Neuroscience*, *104*, 84-97.

# *A few words*

Before I met Nick I had never been to Spain. He found that hard to believe and took me with him as soon as he could. The first time that I went I was absolutely entranced, and could see just why he was so enraptured. For a man who loved fine weather, flowing wine, and flowing conversation, Spain was perfect. He loved the country and the culture, and was very positive about Spanish psychology. He was very pleased that people here are interested in the kind of work that he thought so important. For as long as I knew him he studied Spanish, and as some people will know, he always tried to introduce his talks in Spanish. It was very important to him as a way of showing his respect for his Spanish colleagues, but only they will know how well he succeeded!

Nick particularly loved Barcelona, although learning Catalan may have proved too much for him! He was very grateful for his continuing association with Professor Chamizo, and for her constant generosity to him, and I too am very grateful for her friendship and generosity.

I am sure that I speak for Nick when I say that he would give you his heartfelt wishes that Spanish psychology may continue to flourish.

Leonora Brosan Mackintosh
June 2016

# *Preface. The Construct of Attention and Beyond: Homage to N. J. Mackintosh (1935-2015)*

Ralph R. Miller
Binghamton University, State University of New York, USA

This volume, edited by Professors Trobalon and Chamizo of the University of Barcelona, is dedicated to the memory of Professor N. J. Mackintosh (1935-2015) in honor of his contributions to our understanding of the basic principles of associative learning. Professor Mackintosh received his PhD in Experimental Psychology from Oxford University in 1963 under the supervision of Professor Stuart Sutherland. He went on to serve on the faculties of the University of Oxford, Dalhousie University, the University of Sussex, and finally the University of Cambridge where he served as head of department from 1981 to 2002. At various times he held visiting professorships at the University of Hawaii, Bryn Mawr College, the University of California at Berkeley, Yale University, the University of Pennsylvania, and the University of New South Wales.

Professor Mackintosh's scientific findings and insights were communicated through hundreds of insightful empirical and theoretical articles as well as a number of books, the four of greatest impact being:

1) Sutherland, N. S. and Mackintosh, N. J. (1971). *Mechanisms of Animal Discrimination Learning*.
2) Mackintosh, N. J. (1974). *The Psychology of Animal Learning*.
3) Mackintosh, N. J. (1983). *Conditioning and Associative Learning*.
4) Mackintosh, N. J. (1998; 2011). *IQ and Human Intelligence*.

The most widely cited of these volumes is *The Psychology of Animal Learning*, which is nearly encyclopedic and was reportedly written with minimal notes. Consistent with his monumental memory, in my conversations with

him he recalled aspects of some of my own experiments that I myself had forgotten decades after they had first appeared in print. Notably, the last of Professor Mackintosh's four books reflects a whole second research field in which he also distinguished himself and greatly influenced the assessment of intelligence in humans as well as practices in our schools. However, the present volume focuses on basic learning and cognition, as studied by Professor Mackintosh and the numerous researchers who were inspired by him and his work in this domain.

In addition to his own scientific contributions, an equally significant contribution of his remarkable academic career was his mentorship of innumerable students, postdoctoral fellows, collaborators, and the many visiting scholars who passed through his laboratories at Sussex and Cambridge. Why did so many researchers make the pilgrimage to Sussex and Cambridge to exchange views with Professor Mackintosh? The attraction was not only his depth of knowledge and broad interests, but his readiness to discuss the visitor's research, well-seasoned with his warm hospitality. Here I speak from first-hand experience, having been one of those pilgrims.

A further means through which Professor Mackintosh greatly influenced the study of basic learning processes was in his service as editor-in-chief of both *The Quarterly Journal of Experimental Psychology* and the *Journal of Experimental Psychology: Animal Learning and Behavior*. In his editorial capacity, he advised a whole generation of researchers with his famously constructive decision letters.

In his lifetime, Professor Mackintosh's scientific contributions were recognized through numerous awards including his being elected a Fellow of the Royal Society, and his receiving the Biological Medal and the President's Award from the British Psychological Society. The renown of his empirical and theoretical work gave him the visibility that allowed him to serve as the world's leading spokesman/advocate for the behavioral study of animal learning and cognition.

Professor Mackintosh's own research on learning was far-ranging. But he often returned to the construct of *attention*, which he typically narrowed to *associability* to avoid the excess baggage carried by the construct "attention". His focus on attention started with Sutherland and Mackintosh (1971) in which *selective attention* to attributes of a stimulus was directed by reinforcement and was a conserved quantity. Related to this theme, he published many empirical papers examining attention/associability as assessed in studies of

discrimination in animals. By 1975, his working model (Mackintosh, 1975) had evolved so that attention was no longer rigorously conserved, consistent with variation of arousal influencing the total amount of attention. But attention was still quasi conserved in that increases in attention to cues that best predicted impending biologically significant events were accompanied by decreases in attention to less valid cues that were present on the same training trial. It is worth noting that although Mackintosh invoked selective changes in attention to different cues (or attributes of cues), alternative accounts were proposed by others to address so-called attentional phenomena. In contrast to Mackintosh's mechanism of attention/associability changing as a function of experience, the formal construct of modifiable attention can be circumvented. This is most readily seen in the model of Rescorla and Wagner (1972), in which the associability of a cue is fixed and instead the subject's acquired behavior modulates subsequent learning about the cue. For example, in the visual modality subjects learn what to direct their gaze at, although this approach is more difficult (but not impossible) to apply in some other modalities, such as audition.

Professor Mackintosh's next steps forward in his continuing studies of attention are best seen in McLaren and Mackintosh (2000, 2002; also see McLaren, Kaye, & Mackintosh, 1989). Conventionally, perception and associative learning were viewed as independent sequential processes with perception preceding associative learning. However, the Gestalt psychologists viewed learning as a by-product of the laws of perception. In contrast, McLaren and Mackintosh viewed many perceptual phenomena as by-products of the laws of associative learning. The McLaren and Mackintosh model is a real-time, micro-elemental theory of learning aimed largely at uniting perception with traditional associative learning. The model emphasizes building percepts based on three processes: 1) excitatory within-compound associations between micro-elements that are presented together; 2) inhibitory within-compound associations between micro-elements not presented together but sharing companion micro-elements, and 3) decreased associability of nonreinforced micro-elements scaled to number of presentations along with increased associability of reinforced micro-elements scaled to number of presentations.

Professor Mackintosh's final theoretical statement concerning attention/associability is presented in Pearce and Mackintosh (2010). This model is a hybrid of his 1975 model, which attributed increased attention/associability

to predictive cues of high validity, and the Pearce and Hall (1980) model, which attributed increased associability to cues that are followed by surprising outcomes. As there are phenomena consistent with both mechanisms, Pearce and Mackintosh proposed ways in which these two processes could coexist.

Given the current emphasis in the literature on learning being driven by error reduction, it should be noted that Mackintosh (1975) used a local error reduction rule for learning similar to that of Bush and Mosteller (1951), in which the error that drives associative acquisition for any given cue is the difference between the outcome anticipated based on that cue and the outcome experienced. This contrasts with learning being driven by the total error reduction as assumed in the Rescorla and Wagner (1972) model, in which the error reduced in learning is the difference between the outcome anticipated based on all cues present on a given trial and the outcome experienced. In the Rescorla-Wagner model, total error reduction is the process responsible for cue competition, whereas in Mackintosh (1975) cue competition is produced by decreases in the associability of less valid predictors of the outcome. Later formulations such as Pearce and Mackintosh (2010) used a total error reduction mechanism. But Professor Mackintosh was not strongly committed to total as opposed to local error prediction. He retained the view that apparent total error reduction could be an artifact of changing attention and other processes (see Stout & Miller, 2007, for another account of cue competition that is not predicated on total error reduction). Professor Mackintosh was not theoretically rigid, in that he was willing to entertain a variety of somewhat contradictory models, recognizing that no contemporary model is comprehensive. Rather, he viewed models as heuristic devices to shape our thinking and direct us when designing experiments.

If there was one consistent feature of Professor Mackintosh's theorizing over the years, it was that he maintained a bottom-up orientation to his accounts of learning; that is, he consistently tried to explain so-called "reasoning" through simpler associative processes. More specifically, he viewed dyad associations as the foundation of many cognitive processes that others viewed as instances of top-down reasoning.

I close this preface by briefly describing five projects in my own laboratory that were significantly influenced by illuminating conversations I had with Professor Mackintosh over a fourty-year period. Any foolishness here is

of my own doing, but merit if any is in part a consequence of Professor Mackintosh's insights that greatly influenced my thinking and sometimes my experimental designs.

1) Implications of cue-to-consequence effects for stimulus associability: models that assign an associability to a given cue, whether it is variable as a function of validity or not, are unable to account for the now well documented superiority in learning some cue-outcome associations over other cue-outcome associations. Experiments have demonstrated that these cue-to-consequence effects cannot be explained simply by some cues and some outcomes having greater associability than other cues and outcomes (Garcia & Koelling, 1966; also see Foree & LoLordo, 1973, for an extension of this principle to responses). Seemingly, the only way to capture these effects in conventional models of learning is to posit an associability for the cue-outcome dyad (or cue-response outcome in the instrumental case studied by Foree and LoLordo). This problem challenges all contemporary models of learning (Miller & Grace, 2003).

2) Two quasi-independent types of memory interference: first, presentation of irrelevant stimuli near the time of training, or just prior or during a test, often disrupts acquired behavior, presumably because of competition for the limited capacity of working memory. The magnitude of these disruptions is directly related to the time between the disruptive irrelevant event and target training or testing, i.e., recency effects are observed here. Such interference can be couched in terms of the two memories, target and interfering, interacting because of their common time of activation. A second form of interference, often called associative interference, is often observed when a target association (Cue X-Outcome A) and a potentially interfering association have some but not all elements (e.g., Cue X-Outcome B or Cue Y-Outcome A; see Miller & Escobar, 2002). Our initial work contrasting these two types of interference (Miller, Greco, Marlin, & Balaz, 1985), one dependent on similarity in time and the other on similarity in non-temporal content, respectively, was greatly facilitated by conversations with Professor Mackintosh.

3) Over several decades, my collaborators and I developed a model of acquired behavior that emphasized information processing at the time of test (i.e., the comparator hypothesis, Miller & Matzel, 1988; Denniston, Savastano, & Miller, 2001; Stout & Miller, 2007), in contrast to most other models that focus on information processing during test (e.g., Mackintosh, 1975; Rescorla & Wagner, 1972). This orientation was diametrically opposed to that

taken by Professor Mackintosh in all of his published works, but he generously provided constructive advice to me over the several decades that we developed this model.

4) Professor Mackintosh and I shared reservations concerning whether accounts of learning based on reasoning and inference by the subject ever rose above simply relegating the question to a homunculus. One instance of learning in humans that is frequently explained in terms of inference is causal attribution. This prompted me to seek a demonstration of causal attribution in a nonhuman species, one presumably unlikely to employ complex inferences. Presumably, if rats exhibited behaviors analogous to behaviors that in humans are assumed to reflect causal learning, the same bottom-up associative processes that produced these behaviors in rats might also apply to causal learning in humans. Professor Mackintosh was highly supportive of this project, and in his role as the Devil's Advocate caused us to run additional experiments to assess alternative accounts of our rats' behavior that might have distinguished the behavior of our rats from that of humans exhibiting causal learning (Polack, McConnell, & Miller, 2013).

5) Perception learning: Professor Mackintosh devoted much of his efforts to examining the influence of learning on perception. The keystone to much of his later theorizing about this relationship was Espinet, Iraola, Bennet, and Mackintosh's (1995) demonstration of inhibitory perceptual learning. They found that many interspersed non-reinforced exposures to two compound stimuli, AX and BX, followed by pairings of B with an unconditioned stimulus (US) makes A inhibitory with respect to B or at least to the US. My colleagues and I developed a new technique to differentiate B being inhibitory with respect to A as opposed to the US, and found that indeed B was inhibitory with respect to A (also see Dwyer & Mackintosh, 2002). We also asked whether inhibitory perceptual learning obeys the same rules as conditioned inhibition. In conventional conditioning, few XB- trials followed by B-US trials produces excitatory sensory preconditioning of X; whereas many such trials makes X an conditioned inhibitor (Stout, Miller, & Escobar, 2004). Espinet et al. found that *many* AX/BX trials followed by B-US makes A (and X) inhibitory. We replicated this finding and additionally found that *few* AX/BX followed by B-US makes A (as well as X) excitatory. Hence, perceptual learning and conditioning obey parallel rules at least in this respect (manuscripts in preparation). In both of these projects, conversations with Professor Mackintosh greatly influenced our research.

The following chapters by a number of Professor Mackintosh's most prominent students, postdoctoral fellows and collaborators will give the reader a flavor of both the general focus of his primary interests as well as the diversity of specific questions that he and his collaborators have pursued. Professor Mackintosh's impact on the entire field of basic learning has been enormous.

Thank you, Nick. You will be missed.

## References


Bush, R. R., & Mostellar, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*, 313-323.

Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65-117). Hillsdale, NJ: Erlbaum.

Dwyer, D. M., & Mackintosh, N. J. (2002). Alternating exposure to two compound flavors creates inhibitory associations between their unique features. *Animal Learning & Behavior*, *30*, 201-207.

Espinet, A., Iraola, J. A., Bennett, C. H., & Mackintosh, N. J. (1995). Inhibitory associations between neutral stimuli in flavor-aversion conditioning. *Animal Learning & Behavior*, *23*, 361-368.

Foree, D. D., & LoLordo, V. M. (1973). Attention in the pigeon: The differential effects of food-getting vs. shock avoidance procedures. *Journal of Comparative and Physiological Psychology*, *85*, 551-558.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, *4*, 123-124.

Mackintosh, N. J. (1974). *The psychology of animal learning*. London: Academic Press.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276-298.

Mackintosh, N. J. (1983). *Conditioning and associative learning.* Oxford: Oxford University Press.

Mackintosh, N. J. (1998). *IQ and Human Intelligence*. Oxford: Oxford University Press.

Mackintosh, N. J. (2011). *IQ and Human Intelligence* (2nd ed.). Oxford: Oxford University Press.

McLaren, I. P. L., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. In R. G. M. Morris (Ed.), *Parallel distributed processing: Implications for psychology and neurobiology* (pp. 102-120). Oxford: Clarendon Press / Oxford University Press.

McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, *28*, 211-246.

McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, *30*, 177-200.

Miller, R. R., & Escobar, M. (2002). Associative interference between cues and between outcomes presented together and presented apart: An integration. *Behavioural Processes*, *57*, 163-185.

Miller, R. R., & Grace, R. C. (2003). Conditioning and learning. In A. F. Healy & R. W. Proctor (Eds.), *Experimental psychology*, *Vol 4* (pp. 357-397), of *Handbook of Psychology*, I. B. Weiner (Ed.). New York: John Wiley & Sons.

Miller, R. R., Greco, C., Marlin, N. A., & Balaz, M. A. (1985). Retroactive interference in rats: Independent effects of time and similarity of the interfering event with respect to acquisition. *Quarterly Journal of Experimental Psychology*, *37B*, 81-100.

Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation*, *Vol. 22* (pp. 51-92). San Diego, CA: Academic Press.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552.

Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. In C. Mitchell, & M. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 11-39). Oxford: Oxford University Press.

Polack, C. W., McConnell, B. L., & Miller, R. R. (2013). Associative foundation of causal learning in rats. *Learning & Behavior*, *41*, 25-41.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Stout, S. C., Escobar, M., & Miller, R. R. (2004). Trial number and temporal relationship as joint determinants of second-order conditioning and conditioned inhibition. *Learning & Behavior*, *32*, 230-239.

Stout, S. C., & Miller, R. R. (2007). Sometimes competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, *114*, 759-783.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.