

## SOCIAL LEARNING IN ANIMALS: CATEGORIES AND MECHANISMS

By C. M. HEYES

*Department of Psychology, University College London, Gower Street,  
London WC1E 6BT*

(Received 6 October 1993; revised 7 February 1994; accepted 11 February 1994)

### CONTENTS

I. Introduction . . . . .	207
(1) Overview . . . . .	207
(2) Asocial learning . . . . .	209
(3) Social learning . . . . .	212
II. Stimulus enhancement as single stimulus learning . . . . .	213
(1) Current definitions: local and stimulus enhancement . . . . .	213
(2) Problems . . . . .	214
(a) Local <i>vs.</i> stimulus enhancement . . . . .	214
(b) Stimulus enhancement . . . . .	215
(3) Proposal . . . . .	216
III. Observational conditioning as S–S learning . . . . .	217
(1) Current definition: observational conditioning . . . . .	217
(2) Limitations . . . . .	218
(3) Proposal . . . . .	220
(4) Anomalies . . . . .	221
(a) Selective associability . . . . .	221
(b) Immunization . . . . .	222
(c) Spatial contiguity . . . . .	223
(d) Response evocation <i>vs.</i> learnability . . . . .	223
IV. Observational learning as R–S learning . . . . .	224
(1) Current definitions: imitation, copying and observational learning . . . . .	224
(2) Problems . . . . .	225
(3) Proposal . . . . .	225
(4) Anomalies . . . . .	227
(a) Matching disparate sensory inputs . . . . .	227
(b) Intrinsic reinforcement . . . . .	227
(c) Matched-dependent behaviour . . . . .	228
V. Summary and conclusions . . . . .	229
VI. Acknowledgements . . . . .	229
VII. References . . . . .	229

### I. INTRODUCTION

#### (1) *Overview*

The term social learning refers to learning that is influenced by observation of, or interaction with, another animal (typically a conspecific) or its products (Box, 1984; Galef, 1988). The complementary set is commonly known as ‘individual learning’. However, since in all cases it is ultimately individuals who learn, this usage is

potentially confusing, and the term 'asocial learning' will be used here to refer to learning that does not involve social interaction.

Since formal research on the subject began a century ago (see Galef, 1988 for an historical review), the majority of studies of social learning have either sought to demonstrate that the members of a given species are capable of some form of social learning, or examined the role of social learning in the ontogeny of adaptive behaviour. Relatively little attempt has been made to investigate the mechanisms, rather than the functions, of social learning (Davis, 1973; Galef, 1988). In contrast, studies of asocial learning have largely eschewed questions about adaptive function, and have instead provided a substantial body of information, 'animal learning theory', relating to the psychological mechanisms of learning. Investigators of social learning seldom refer to animal learning theory, even when they are discussing mechanisms. This implies that animal learning theory is not applicable to social learning; that different mechanisms are responsible for social and asocial learning. On the other hand, in formulating general principles of learning, investigators of asocial learning rarely state that social learning may be an exception. It would appear that many of these latter researchers assume that the same mechanisms are responsible for social and asocial learning.

To what extent are the mechanisms of social learning the same as those of asocial learning? The purpose of the present review is to facilitate research addressing this question; the application of animal learning theory to social learning. This aim is pursued by relating the categories of learning used by students of asocial learning to those used by investigators of social learning. The categories used for asocial learning do not map onto those used for social learning in a simple, one-one fashion and, at present, there is no 'wiring diagram' relating the two sets. Consequently, it is not clear which social and asocial learning phenomena, if any, occur under similar conditions, and, since the conditions of learning are indicators of its mechanisms, it is difficult even to address the question of whether the same or similar mechanisms are responsible.

In the remainder of Part I, an overview is given of the categorization schemes currently applied to asocial and social learning, respectively, and some of the weaknesses of the latter scheme are noted. The categories of learning used by investigators of social and asocial learning, respectively, are the products of conceptual evolution rather than design. They are based on a combination of terms and distinctions that were of use to practising scientists at various times and for various purposes. They do not derive from the kind of formal principles or criteria of classification found in, for example, zoological systematics. However, it is possible to summarize the most important distinctions used by contemporary investigators of asocial and social learning, and this has been achieved recently in Rescorla's (1988) 'framework for the study of learning', and Galef's (1988) glossary of terms used in research on imitation. These are the sources of the account of classification practices given in the remainder of this section.

In Parts II-IV, an attempt is made to subsume social learning phenomena within the categorization scheme commonly used by investigators of asocial learning. Within each Part, the current definition of a category of social learning is given, and the weaknesses of that definition are discussed. An alternative definition is then proposed, one which equates the target category of social learning with an existing category of asocial learning. This cross-classification exercise indicates that most, if not all, social learning

Table 1. *Categories of learning*

Experience	Behaviour change	Category
Stimulus	Response evocation	Habituation Sensitization
	Learning	Latent inhibition Perceptual learning
	Response evocation	Pavlovian conditioning
Stimulus-stimulus	Learnability	Blocking Overshadowing
	Response evocation	Instrumental conditioning
Response-reinforcer	Learnability	Blocking Overshadowing

phenomena resemble asocial learning phenomena in terms of their conditions of occurrence, and thereby provides *prima facie* evidence that the two sets of phenomena are mediated by common mechanisms. The combined classification also indicates, through its use of key terms, where information about common mechanisms may be found in the literature on asocial learning, and raises the possibility that there are a number of functionally relevant social learning phenomena which have not yet been identified.

Showing that social learning phenomena can be subsumed within a categorization scheme used for asocial learning, would, at most, encourage the hypothesis that similar mechanisms are responsible for social and asocial learning, and facilitate research investigating that hypothesis. It certainly would not either imply that contemporary animal learning theory provides a sufficient account of the mechanisms of social learning, or act as a substitute for empirical research. Consequently, Parts III and IV include discussion of recent empirical work which apparently shows that certain social learning phenomena involve mechanisms distinct from those normally mediating asocial learning. These studies concern potential anomalies with respect to a unified account of the mechanisms of social and asocial learning.

## (2) *Asocial learning*

Within contemporary learning theory, learning is regarded as change in an animal that is caused by a specific experience at a certain time,  $t_1$ , and that is detectable later,  $t_2$ , in the animal's behaviour (Rescorla, 1988). As illustrated in Table 1, investigators of asocial learning distinguish categories of learning according to the type of experience at  $t_1$  which causes a change in the organism, and the type of behaviour in which this change is detected at  $t_2$  (Rescorla, 1988).

There are three, commonly recognized, types of experience: (i) a single stimulus (an object or event in the animal's environment) (S), (ii) a relationship between two stimuli (S-S), and (iii) a relationship between some action or response made by the animal, and a stimulus (R-S). (In the latter case, the stimulus is known as a 'reinforcer'.) For example, a frog sitting among some reeds might experience, on one occasion or repeatedly, a rustling sound (S), a rustling sound accompanied by the appearance of a predator (S-S), or access to a prey item when it hops from one place to another (R-S).

Each of these types of experience can result in changes in behaviour of various kinds. Taking experience of a single stimulus as an example, hearing a rustling sound may result in a frog becoming more or less likely to respond to subsequent experiences of that sound by turning its head. If the frog became more likely to respond by turning its head, or in some other way, it would be said to have shown *sensitization*, and if it became less likely to respond, the frog would be said to have shown *habituation*. In both cases the effects of single stimulus experience would have been manifested in a change in response evocation.

Alternatively, the effect on the frog of hearing the rustling sound alone might be detected at  $t_2$  in the 'learnability' of a relationship between that stimulus and another event, for example, in the frog's capacity to learn about a relationship between the rustling sound and the arrival of a predator. If prior, single stimulus experience interfered with the frog's capacity to learn about this relationship, or any other relationship between the rustling sound and a second stimulus, the effect on the frog of the single stimulus experience would be an example of *latent inhibition*. If the single stimulus experience facilitated subsequent learning about such relationships, it would be an example of *perceptual learning*.

A change in an animal resulting from exposure to a relationship between two stimuli is an example of *Pavlovian conditioning* (or classical conditioning), and a change resulting from experience of a response–reinforcer relationship is an example of *instrumental learning* (or operant conditioning), regardless of the manner in which it is detected in the animal's behaviour. However, like the effects of single stimulus experience, those of exposure to relationships between events may be detected through, *inter alia*, changes in response evocation and changes in learnability; in their capacity to modulate the effects of exposure to other relationships.

Changes in response evocation tend to vary with the nature of the relationship between the two events (positive or negative), and the value of the second event for the animal (appetitive or aversive). The relationship is positive if the first event predicts the second event, and it is negative if the occurrence of the first event predicts that the second event will not occur. For example, when the relationship between the two events is positive (*excitatory conditioning*), and the second event is something attractive, such as the arrival of food, there tends to be an increase in response evocation. When the relationship is between two stimuli, the first stimulus becomes more likely to evoke a response, and when the relationship is between a response and a reinforcer, the response becomes more intense or likely to occur in future. However, when the relationship between the two events is positive, and the second event is something aversive, such as the arrival of a predator, the effect of exposure to the relationship is usually detected through a decline in response evocation. In the case of a stimulus–stimulus relationship, the first event becomes less likely to elicit an active response, and in the case of a response–reinforcer relationship, the response is less likely to occur.

Similar tendencies exist with respect to negative relationships. When the second event is negatively correlated with the first (*inhibitory conditioning*), and the second event is appetitive, there tends to be a decline in response evocation. In the case of stimulus–stimulus relationships, there is a decline in responding to the first stimulus, and in the case of response–reinforcer relationships, the response is less likely to occur.

Finally, when there is a negative relationship and the second event is aversive, an increase in response evocation is often detected. For stimulus–stimulus relationships, responding to the first stimulus increases, and for response–reinforcer relationships, the probability or intensity of the response increases.

The effects of exposure to stimulus–stimulus and response–reinforcer relationships are detected, not only through changes in response evocation, but also through changes in the learnability of other relationships. Thus, the effect of experience of a relationship between  $S_1$  and  $S_2$  may become apparent, not through a change in the degree to which  $S_1$  evokes a response, but through a decline in responding to another stimulus,  $S_3$ , that was paired with  $S_2$  at around the same time as  $S_1$ . To illustrate, consider three frogs, A, B and C. On several occasions, frog A sees ripples ( $S_3$ ) appear in a nearby pool just before the arrival of a predator ( $S_3$ – $S_2$ ), while frog B sees ripples ( $S_3$ ) and simultaneously hears the sound of rustling in the reeds ( $S_1$ ) just before the arrival of a predator ( $S_3$  +  $S_1$ – $S_2$ ). Frog C has the same experience as frog B, but C had previously heard rustling ( $S_1$ ) alone prior to the appearance of a predator ( $S_1$ – $S_2$ ,  $S_3$  +  $S_1$ – $S_2$ ). When each of the frogs subsequently sees ripples in the pool ( $S_3$ ), it is likely that frog B will respond less vigorously than frog A, and that frog C will respond less vigorously than frog B. The contrast between frog A and frog B in responding to the ripples ( $S_3$ ) would be described as an example of *overshadowing*; experience of the  $S_1$ – $S_2$  at the same time as the  $S_3$ – $S_2$  relationship reduced the degree to which B learned about the  $S_3$ – $S_2$  relationship. The contrast between frog B and frog C, on the other hand, would be described as an example of *blocking*; prior experience of the  $S_1$ – $S_2$  relationship reduced the degree to which C learned about the  $S_3$ – $S_2$  relationship when it was experienced simultaneously with the  $S_1$ – $S_2$  relationship. In both overshadowing and blocking, the effects of exposure to one stimulus–stimulus or response–reinforcer relationship is detected through attenuation of the effects of exposure to another such relationship.

In summary, asocial learning has been categorized with respect to the type of experience that gives rise to a change in the animal, and the type of behaviour change in which it is detected. The three main types of experience are of a single stimulus, a stimulus–stimulus (S–S) relationship, and a response–reinforcer (R–S) relationship. Experience of event relationships is further differentiated according to whether the events are positively or negatively correlated, and whether the second event is appetitive or aversive. Each type of experience may be detected through changes in response evocation and learnability.

As the foregoing discussion indicates, investigators of asocial learning distinguish categories of learning with reference to the observable conditions of its occurrence, not to the mechanisms thought to be responsible for learning. Thus, when dealing with a particular example of learning, in principle there are two distinct questions to be addressed: What were the conditions of learning (and therefore what is the category of learning to which this example belongs), and what is the mechanism responsible for this example of learning? In practice, however, members of the same category of learning, that is, specimens of learning which occur under similar conditions, have been found to have the same underlying mechanisms.

A learning mechanism is understood within most contemporary animal learning theory to be something inside an individual, the arrangement and action of which is responsible for a specific experience resulting in a change in behaviour. Such

mechanisms are assumed to be components of the central nervous system, but they have been investigated primarily using the methods and theoretical resources of psychology, rather than neuroscience, and are therefore characterized in functional terms. For example, an associative learning mechanism is one that produces, under specified conditions, increments and decrements in the strength of a connection between psychological representations (see Dickinson, 1980 or Domjan, 1993 for a synopsis of the mechanisms postulated by contemporary animal learning theory).

### (3) *Social learning*

Compared with the classification of asocial learning, that of social learning has been somewhat rudimentary (Galef, 1988). There are relatively few discrete categories, they do not form a hierarchy, it is rarely possible to assign an example of social learning to one category rather than another with certainty, and there is little consensus regarding the use of category labels. One of the most important differences between the asocial and social learning classification schemes is that the latter does not distinguish types of learning according to the way in which learning becomes apparent in behaviour. Throughout its history, the study of social learning has been dominated by an interest in the degree to which animals can imitate and/or transmit behaviour, one to another (Galef, 1988; Heyes, in press; Whiten & Ham, 1993). Consequently, little attention has been paid to the possibility that social learning may have a variety of effects in addition to the production by the learner, or 'observer', of behaviour resembling that of an observed animal, or 'demonstrator'. Thus, matching behaviour, behaviour like that of a demonstrator, is the only widely recognized outcome of social learning.

Varieties of social learning are distinguished according to the role of the demonstrator in generating matching behaviour on the part of the observer. Thus, the demonstrator's behaviour or its products (e.g. scent marks, excavations) may (i) increase the probability that the observer will attend to the stimuli with which it, the demonstrator, interacts (*local enhancement*), (ii) increase the probability that the observer will interact with stimuli of the same physical type as those with which the demonstrator interacts (*stimulus enhancement*), or act as (iii) an unconditioned stimulus eliciting a matching response (*observational conditioning*), (iv) a discriminative stimulus (*matched-dependent behaviour*), or (v) a model within a goal-directed (*imitation*) or non-goal-directed (*copying*) process (Galef, 1988; Whiten & Ham, 1993).

Each of these categories will be considered further below, but several features of the demonstrator–role principle of classification should be noted in anticipation of more detailed discussion. First, unlike type of experience, which is used to classify asocial learning, demonstrator–role cannot be directly observed and manipulated by an investigator. Both type of experience and demonstrator–role are theory-laden bases of classification, in that their application is guided by hypotheses concerning the mechanisms responsible for learning. However, while the asocial learning scheme uses a directly observable condition of learning to distinguish types, and treats type as an indicator of underlying mechanism, the social learning scheme uses type of mechanism as a basis for classification in its own right.

Second, the demonstrator–role principle of classification appeals to mechanisms that were postulated by scientists working in various research traditions at various times, and most of these mechanisms have never been satisfactorily elucidated (Davis, 1973;

Galef, 1976, 1988). For example, Thorpe (1956), an ethologist and naturalist, defined local enhancement with reference to 'attention', and imitation with reference to 'self-consciousness' and goal-directedness, but neither he nor subsequent users of his terminology have indicated how attention differs from perception, how goal-directed processes differ from those that are not goal-directed, or what was meant by 'self-consciousness'. On the other hand, the term 'matched-dependent behaviour' was coined by Miller & Dollard (1941), who were comparative psychologists working within the tradition of Hullian behaviourism. They gave a detailed account of the role of the demonstrator, as a discriminative stimulus, in this kind of social learning, but their account does not describe a 'mechanism' in the contemporary sense (see above). As neobehaviourists, Millar and Dollard did not attempt to explain learning with reference to internal states and operations.

Thus, via the concept of demonstrator-role, social learning is classified on the basis of ill-specified, unobservable mechanisms. One of the consequences of this practice is that it is rarely possible to determine with certainty the class to which a social learning phenomenon belongs. Under any circumstances, demonstrator roles would be difficult to distinguish empirically, and, exacerbating the problem, most examples of social learning derive from field and laboratory studies in which role-relevant variables have not been manipulated. Furthermore, many categories of social learning are not mutually exclusive because it is possible in principle for a demonstrator to play several roles simultaneously. For example, through a single action, a demonstrator might attract an observer's attention to a stimulus (local enhancement), increase the probability that the observer will be exposed to stimuli of the same physical type as those with which the demonstrator is interacting (stimulus enhancement), and either provide an unconditioned stimulus eliciting a matching response from the observer (observational conditioning), act as a discriminative stimulus (matched-dependent behaviour), or act as a model (copying or imitation).

## II. STIMULUS ENHANCEMENT AS SINGLE STIMULUS LEARNING

### (1) *Current definitions: local and stimulus enhancement*

An example of social learning is categorized as *local enhancement* (Thorpe, 1956) if exposure to the demonstrator animal or its products (e.g. scent cues, excavations) draws the observer's attention to the stimuli with which the demonstrator was interacting. For instance, rats (*Rattus norvegicus*) are more likely to consume novel foods that have been scent marked by conspecifics than other novel foods in the same area (Galef & Beck, 1985). If this is because the presence of scent marks increases the attention paid by observer rats to the food eaten by the demonstrator, then it is an example of local enhancement.

*Stimulus enhancement* (Spence, 1937) differs from local enhancement in two respects: the demonstrator's behaviour (i) increases the probability that the observer will be exposed, rather than attend, to certain stimuli, and (ii) the stimuli in question are not only those tokens with which the demonstrator interacts, but all token stimuli of the same physical type. For example, in some experiments using the 'duplicate cage method', pairs of monkeys (Warden & Jackson, 1935) were placed in adjacent cages containing identical sets of objects, and on each trial one of the monkeys, the

demonstrator, was allowed to select an object from its set while being observed by its partner. When the observer monkey was then given access to the objects in its own cage, it tended to select one of the same physical type as that chosen by the demonstrator. If this was because the demonstrator's action increased the probability that the observer would be exposed to stimuli of a certain physical type, then it is an example of stimulus enhancement.

## (2) *Problems*

### (a) *Local vs. stimulus enhancement*

Both of the criteria used to differentiate local and stimulus enhancement are problematic. First, it is difficult to distinguish empirically cases in which a demonstrator increases exposure rather than attention to stimuli, and, partly in consequence, it is not clear what investigators of social learning take to be the difference between the two at a conceptual level. In an attempt to overcome the empirical problem, one might classify as effects on stimulus exposure (and therefore as examples of stimulus enhancement) cases in which the demonstrator's activity apparently increases the probability that a stimulus will enter the perceptual range of the observer, and as effects on stimulus attention (local enhancement) hypothetical cases in which the demonstrator's activity increases the probability that the observer will orient towards particular stimuli that are already within range. However, this is not a satisfactory solution because orienting responses (e.g. direction of gaze) are difficult to measure reliably (LoLordo & Ross, 1990), and there is not a single example of social learning for which data on orienting responses are available. Furthermore, in duplicate cage experiments, conspecific observation apparently affects orientation to stimuli that are within perceptual range, and yet they have been cited as potential examples of stimulus, rather than local, enhancement (Galef, 1988), i.e. as effects of a demonstrator on exposure rather than attention. This implies that, whatever is understood to be the difference between exposure and attention, and therefore between stimulus and local enhancement, it cannot be operationalized with reference to the distinction between bringing stimuli into perceptual range, and directing orienting responses to stimuli that are already within range.

The second criterion specifies that in local enhancement the demonstrator affects the observer's behaviour toward only those stimuli with which the demonstrator interacts, while in stimulus enhancement it affects the observer's behaviour towards all stimuli with the same physical characteristics as those contacted by the demonstrator. This is problematic because it implies either that local enhancement is not a form of learning, or that local enhancement and stimulus enhancement cannot be distinguished in terms of demonstrator role. Local enhancement would not conventionally be regarded as a form of learning if it comprised cases in which the demonstrator or its products affect the observer's behaviour only when they are physically present. For example, if one found that rats preferentially consume scent marked foods, but do not subsequently discriminate among foods that they have and have not encountered with scent marks, and if one classified this as an example of local enhancement, then local enhancement would not be a form of learning.

Alternatively, one might classify as local enhancement only those cases in which, when all traces of the demonstrator have gone, the observer continues to behave



differentially towards the stimuli with which the demonstrator interacted. However, when local enhancement is circumscribed in this way, the difference between local and stimulus enhancement is purely procedural; it is a matter of what an investigator has measured, rather than of what a demonstrator animal has done, and is therefore anomalous with respect to the usual method of distinguishing types of social learning. Consider, for example, a rat that has a choice of two food items  $A_1$  and  $B_1$ , and that consumes more of item  $A_1$  because it has been scent marked by a conspecific. The rat could be said to have shown local enhancement if it continued to behave differentially towards  $A_1$  and  $B_1$  when the scent marks had gone. However, if the rat behaved differentially towards  $A_1$  and  $B_1$  in the absence of scent marks, then surely, if given the chance, it would provide evidence of stimulus enhancement by behaving differentially towards two other food items,  $A_2$  and  $B_2$ , that have exactly the same physical characteristics as  $A_1$  and  $B_1$ . Since the two pairs of stimuli have exactly the same physical characteristics, the rat could not tell them apart, and therefore could not be expected to show a bias in relation to one of the pairs of items and not the other.

Thus, it would appear that local enhancement and stimulus enhancement cannot be distinguished satisfactorily using the methodological and conceptual tools traditionally applied to the task by investigators of social learning. It may be possible to tease them apart using data on orienting responses and with the understanding that they refer, not to different mechanisms, but to different ways in which the effect of a demonstrator on an observer may be measured. However, it is not clear that it is worth labouring to fortify or reconstruct the distinction between local and stimulus enhancement, because it has no apparent relevance either to the adaptive function or, in its reconstructed form, the mechanisms of social learning. Instead, the distinction between local and stimulus enhancement seems to be something of an historical accident, resulting from the fact that two investigators from different research traditions, Thorpe (1956, local enhancement) and Spence (1937, stimulus enhancement), expressed similar ideas in rather different terms.

(b) *Stimulus enhancement*

Galef (1988) has suggested that, in view of the problems associated with the distinction between local and stimulus enhancement, the category of local enhancement should be discarded, and the term stimulus enhancement used to refer to all cases in which a demonstrator increases the probability of matching behaviour on the part of an observer by enhancing the observer's exposure to stimuli of the same physical type as those with which the demonstrator interacts. This would indeed circumvent the problems considered above, but as long as stimulus enhancement is defined with reference to the production of matching behaviour, two substantial problems will remain.

First, stimulus enhancement is not an exclusive category. For example, rhesus monkeys that have observed a conspecific behaving fearfully in the presence of snakes subsequently also behave fearfully in response to snake stimuli (Cook & Mineka, 1988). In this case, the behaviour of the demonstrator may facilitate matching behaviour on the part of the observer by both enhancing the observer's exposure to snake stimuli (stimulus enhancement) and acting as an unconditioned stimulus for fear on the part of the observer (observational conditioning).

Second, even according to Galef's (1988) definition, stimulus enhancement is a heterogeneous category that combines cases in which enhanced stimulus exposure is sufficient to produce matching behaviour, and those in which it will yield matching behaviour only if the observer is subsequently and independently exposed to certain types of experience. In the example, mentioned above, of rats that acquire a preference for a food that has been scent marked by a conspecific (Galef & Beck, 1985), enhanced stimulus exposure may be sufficient to produce matching behaviour. That is, after encountering scent marked food, rats may show dietary preferences like those of the marker/demonstrator, even when the marked food is no more nutritious or enjoyable to eat than diets with other flavours and/or odours. However, in putative cases of stimulus enhancement, the role of enhanced stimulus exposure is far from sufficient. For example, when thirsty rats are given access to a lever which, when depressed, causes the delivery of water, those that have previously observed a conspecific pressing the lever, learn to do so in fewer trials than rats that have not observed a conspecific pressing the lever (Zentall & Levine, 1972). In this case, it is likely that the demonstrator facilitates matching behaviour by enhancing the observer's exposure to the lever, but that the observer would not press the lever reliably if that action was not followed by the delivery of food.

Given this heterogeneity within the category of social enhancement, its use cannot be defended on the grounds that it indicates something about the functional significance of a social learning phenomenon. When socially enhanced stimulus exposure is sufficient to produce matching behaviour, its occurrence under free-living conditions is likely to support social transmission of behaviour. However, if the production of matching behaviour depends, in addition, on the observer independently encountering certain event relationships, the inference from social enhancement to social transmission is much more uncertain.

Thus, even if Galef's (1988) policy with respect to social enhancement were implemented, a single phenomenon could be classified as both social enhancement and some other type of social learning, and any pair of phenomena classified as social enhancement could vary widely in the extent of the demonstrator's contribution to the production of matching behaviour. Consequently, identification of a social learning phenomenon as an example of stimulus enhancement would provide little information about either the conditions of its occurrence, the mechanism underlying the phenomenon, or its functional significance.

### (3) *Proposal*

It is proposed that stimulus enhancement be regarded as a subset of single stimulus learning, as the latter is currently defined by investigators of social learning. Thus, stimulus enhancement occurs when observation of a demonstrator (or its products) exposes the observer to a single stimulus (rather than a relationship between events) at  $t_1$ , and single stimulus exposure effects a change in the observer detected, in any behaviour, at  $t_2$ .

Following Galef (1988), this proposal does not distinguish local from stimulus enhancement and therefore does not encounter the difficulties associated with that distinction. Its implementation would not involve a radical change in usage of the term stimulus enhancement, since it has always been understood to refer to social learning

phenomena in which a demonstrator enhances single stimulus exposure. However, if stimulus enhancement were aligned with single stimulus learning in the proposed manner, it would become one of several mutually exclusive categories of social learning (discussed below), and the problem of internal heterogeneity could be readily overcome. Stimulus enhancement would no longer be defined with reference to matching behaviour, and therefore variations in the degree to which stimulus exposure is sufficient to produce this effect would no longer be a classification issue. Instead, one would ask how the change induced by stimulus exposure had been detected, and subclassify examples of social enhancement accordingly. Thus, if demonstrator-mediated stimulus exposure was detected in a reduction in responsivity to the stimulus it would be a case of *habituation*, and if it was detected in an increase in responsivity it would be *sensitization*. *Latent inhibition* would occur if the effect of demonstrator-mediated stimulus exposure were detected through a reduction in the rate at which the observer subsequently learned about a relationship between the exposed stimulus and another stimulus, and if demonstrator-mediated stimulus exposure enhanced the rate of such learning it would be an example of *perceptual learning*.

The word 'social' might be used as a prefix with each of these terms to indicate that stimulus exposure was enhanced through social interaction. However, at present, there is no reason to suppose that the mechanisms responsible for habituation, sensitization, latent inhibition and perceptual learning are different when stimulus exposure is and is not socially enhanced. This is an empirical issue, and the principal purpose of the proposed definition of social enhancement is to stimulate research investigating this question.

A further potential advantage of the proposed definition of social enhancement is that it draws attention to the possibility that socially enhanced stimulus exposure may be, not only insufficient to yield matching behaviour, but equally likely to increase and to decrease the probability of matching behaviour. For example, as a result of socially enhanced exposure to a stimulus, an animal may become less responsive to that stimulus (habituation), or less likely to learn about a relationship involving that stimulus (latent inhibition). If the behaviour through which the demonstrator enhanced stimulus exposure was either a response to the enhanced stimulus, or based on learning about a relationship involving that stimulus, this would presumably reduce the probability of matching behaviour.

Only empirical work, seeking and examining examples of social habituation and other stimulus enhancement phenomena, will reveal the true distribution of cases in which it is and is not likely to yield matching behaviour and therefore provide a basis for social transmission. However, given its potential to reduce the probability of matching behaviour, it seems strange that stimulus enhancement has ever been regarded as a 'mimetic process' (Whiten & Ham, 1993); one that has evolved because it promotes social transmission.

### III. OBSERVATIONAL CONDITIONING AS S-S LEARNING

#### (1) *Current definition: observational conditioning*

Unlike local enhancement and stimulus enhancement, use of the term observational conditioning represents an attempt to integrate research on social and asocial learning. It was introduced by Cook, Mineka, Wolkenstein & Laitsch (1985), and is understood

to be Pavlovian conditioning (stimulus–stimulus learning) in which an unconditioned response on the part of a demonstrator acts as an unconditioned stimulus (S<sub>2</sub>) eliciting a matching response on the part of the observer. The observer experiences this S<sub>2</sub> with another stimulus, the one to which the demonstrator is responding (S<sub>1</sub>), and as a result of exposure to this relationship, subsequently makes the same response to the S<sub>1</sub> as did the demonstrator (Galef, 1988; Whiten & Ham, 1993). Rhesus monkeys (*Macaca mulatta*) can acquire fear of snakes through observational conditioning (Mineka, Davidson, Cook & Keir, 1984; Cook *et al.*, 1985). In several experiments, laboratory-reared monkeys were allowed to observe a wild-born conspecific behaving fearfully (S<sub>2</sub>) when confronted with a snake (S<sub>1</sub>). Prior to this experience, the observers did not show any fear of snakes, but they behaved in an agitated manner during observation, and when subsequently presented with snake stimuli in the absence of the demonstrator, they displayed fear. Other examples of observational conditioning involve attack behaviour in European blackbirds (*Turdus merula*), Australian zebra finches (*Taeniopygia guttata*) (Curio, 1988), and octopuses (*Octopus vulgaris*) (Fiorito & Scotto, 1992).

### (2) *Limitations*

Compared with local enhancement and stimulus enhancement, observational conditioning is a clearly and coherently defined category of social learning. However, the way in which observational conditioning is currently circumscribed is inappropriately restrictive in a number of respects.

First, there is no evidence that an observer must make a response (matching or otherwise) to the S<sub>2</sub>, the demonstrator's behaviour, in order to learn about the relationship between S<sub>1</sub> and S<sub>2</sub>. Cook, Mineka and their associates (Mineka *et al.*, 1984; Cook *et al.*, 1985) have found a positive correlation between the amount of fear exhibited by observer monkeys during observation and when subsequently confronted with snake stimuli, but fear during observation could be a consequence of learning about the snake–demonstrator fear relationship, rather than a necessary condition for that learning. To find out whether in this case observational conditioning requires the experience or expression of fear, one could, for example, find out whether the effect occurs when observers are treated with an anxiolytic drug prior to observation.

Second, even if it were necessary for the observer to respond to the demonstrator's behaviour during observation, there is no reason to assume that, in order for learning about the S<sub>1</sub>–S<sub>2</sub> relationship to occur, the observer's response to the demonstrator's behaviour must be unconditioned in the sense that it would occur regardless of the observer's prior experience of the S<sub>2</sub>. For example, monkeys may respond fearfully to the fearful behaviour of a conspecific regardless of their prior experience, or only when conspecific fear has been related, in the monkeys' experience, with other, fear-provoking stimuli, such as the arrival of a predator. In the latter case, one might apply a convention used in the literature on asocial learning, and describe the phenomenon as *higher order* observational conditioning (the S<sub>2</sub> in the current episode of S–S learning acquired its potential to fulfil that role by acting as the S<sub>1</sub> in a prior episode of S–S learning). However, it is an empirical question whether observational conditioning is more or less likely to occur when the observer's response to the demonstrator's behaviour is itself conditioned.

Third, and more generally, it is not clear why the term observational conditioning should be reserved for cases in which a demonstrator mediates exposure to an S-S relationship by acting as a source of the second stimulus. There are numerous examples of social learning in which, although the demonstrator's behaviour does not constitute the second stimulus, its activity appears to increase the probability that an observer will be exposed to a stimulus-stimulus relationship. These phenomena cannot be readily classified within the scheme currently used by investigators of social learning, and, since they relate to the observer's experience of a stimulus-stimulus relationship, 'observational conditioning' would appear to be their proper description.

One such example, in which the demonstrator is not the source of the S<sub>2</sub>, comes from an experiment in which pigeons (*Columba livia*) were each presented with a box of grain covered with a sheet of white paper and marked with a red spot (Palameta & Lefebvre, 1985). Birds that had observed a conspecific piercing the red part of the paper (S<sub>1</sub>) and eating the grain inside the box (S<sub>2</sub>), pierced the paper on their own box more quickly and with fewer pecks than did birds that had observed a demonstrator either piercing the cover of an empty box (S<sub>1</sub> only), or eating grain through a preexisting hole in the cover (S<sub>2</sub> only).

The foregoing example also draws attention to a fourth respect in which the current definition of observational conditioning is too restrictive. Contrary to the view of some authors (e.g. Whiten & Ham, 1993), there is evidence that a demonstrator can mediate stimulus-stimulus learning, not only when the second stimulus is aversive, but also when it is appetitive. In Palameta & Lefebvre's (1985) experiment, the behaviour of the piercing-and-eating demonstrators served to expose their observers to a relationship between the red spot (S<sub>1</sub>) and an appetitive stimulus, food (S<sub>2</sub>).

Fifth, as it is currently defined, observational conditioning does not explicitly include cases in which a demonstrator mediates exposure to a negative relationship between stimuli; one in which the occurrence of the S<sub>1</sub> predicts that the S<sub>2</sub> will not occur. An example of this kind of observational conditioning was provided by rhesus monkeys in an experiment on object discrimination (Darby & Riopelle, 1959). On each trial in this experiment, an observer monkey saw a conspecific demonstrator displace one of two distinctive, novel objects which were covering food wells. On some trials, there was food in the well under the selected object, which the demonstrator was allowed to eat, and on other trials the demonstrator's choice was not rewarded. After observing the demonstrator's choice, the observer was presented with the same pair of objects. The results showed that when the demonstrator's choice was rewarded, the observer tended to displace the same object as the demonstrator, and when it was not rewarded, the observer tended to choose the other object. The latter finding suggests that, on non-rewarded trials, the observers learned a negative relationship between the stimulus displaced by the demonstrator (S<sub>1</sub>) and food (S<sub>2</sub>).

Finally, the current definition of observational conditioning is too narrow because it includes only those cases of S-S learning by observation which yield matching behaviour. Socially-mediated exposure to a relationship among stimuli does not inevitably, or perhaps even typically, lead to matching behaviour. Darby & Riopelle's (1959) experiment provided one example in which such experience reduced the probability of matching behaviour. Another example was provided by an experiment on zebra finches (*Taeniopygia guttata*) in which the effects on observers of socially

mediated exposure to a relationship between  $S_1$  and  $S_2$  was detected in the extent to which they learned about a relationship between  $S_3$  and  $S_2$  (Beauchamp & Kacelnik, 1991). In this study, naive birds were placed in an operant chamber with a conspecific demonstrator. Periodically, a light went on ( $S_3$ ), and food ( $S_2$ ) was delivered to the birds via a food tray. Half of the observers were accompanied by 'knowledgeable' demonstrators, that is, birds that approached the food tray as soon as they saw the light go on. These observers were therefore exposed to a second relationship, between the demonstrator approaching the tray ( $S_1$ ) and the arrival of food ( $S_2$ ). The other observers had 'non-knowledgeable' demonstrators that were, on average, no more likely to approach the tray just before food delivery than at any other time. Beauchamp & Kacelnik (1991) found that, both when they were tested alone and with their demonstrators, observers of knowledgeable demonstrators were *less* likely than observers of non-knowledgeable demonstrators to approach the tray when the light went on. That is, the birds that had been exposed, by their demonstrators, to the  $S_1$ - $S_2$  relationship, exhibited less matching behaviour than the birds that had not been exposed to a relationship between stimuli by their demonstrators.

In summary, the current definition of observational conditioning is too narrow because it implies, in the face of argument and evidence to the contrary, that socially-mediated exposure to a relationship among stimuli will affect an observer's behaviour only if the demonstrator's behaviour constitutes an aversive unconditioned stimulus to which the observer responds during observation, and that the effect is detectable only in the production of matching behaviour.

### (3) *Proposal*

It is proposed that observational conditioning be regarded as a subset of stimulus-stimulus learning in which observation of a demonstrator exposes the observer to a relationship between stimuli at  $t_1$ , and exposure to this relationship effects a change in the observer detected, in any behaviour, at  $t_2$ . Examples of observational conditioning, thus defined, could be subdivided into four groups: *excitatory-appetitive* (positive  $S_1$ - $S_2$  relationship, appetitive  $S_2$ ) (e.g. Palameta & Lefebvre, 1985, see above); *excitatory-aversive* (positive  $S_1$ - $S_2$  relationship, aversive  $S_2$ ) (e.g. Cook *et al.*, 1985, see above); *inhibitory-appetitive* (negative  $S_1$ - $S_2$  relationship, appetitive  $S_2$ ) (e.g. Darby & Riopelle, 1959, see above); *inhibitory-aversive* (negative  $S_1$ - $S_2$  relationship, aversive  $S_2$ ). Lore, Blanc & Suedfeld (1971) provided a potential example of inhibitory-aversive observational conditioning by showing that rats which had observed a conspecific sniffing a candle flame and being burned subsequently avoided contact with the candle.

Observational conditioning phenomena could be further differentiated according to whether the effects of socially-mediated exposure to a stimulus-stimulus relationship are detected through a change in the observer's response to the  $S_1$  or to another stimulus that has been paired with the  $S_2$ . Observational conditioning of the latter kind (e.g. Beauchamp & Kacelnik, 1991, see above) may be least likely to promote matching behaviour and therefore to support the social transmission of behaviour when it occurs under free-living conditions. However, there is no apparent reason to suppose that observational conditioning generally promotes such transmission.

The current proposal is to broaden usage of the term observational conditioning. The

first fully documented and carefully analysed example of observational conditioning was of the excitatory-aversive type (e.g. Cook *et al.*, 1985), but, the proposal assumes, this historical contingency is insufficient reason to exclude from the category of observational conditioning other phenomena in which exposure to a relationship among stimuli is socially mediated.

#### (4) *Anomalies*

It could be argued that, while the historical contingency may be unimportant, there is another reason to resist assimilating observational conditioning into a unified taxonomy of social and asocial learning. Such assimilation may be taken to imply that observational conditioning does not differ in any significant way from other, asocial examples of stimulus-stimulus learning, and yet there are studies purporting to show that observational conditioning is anomalous; that it differs from most examples of asocial, stimulus-stimulus learning. The classification of social learning proposed in this article is intended to facilitate research on the mechanisms of social learning, not to embody firm conclusions about those mechanisms. However, the apparently anomalous features of observational conditioning will be discussed briefly because it might be counterproductive to align observational conditioning with other, asocial learning phenomena if it were already clear that distinctive mechanisms are responsible.

##### (a) *Selective associability*

In several experiments, Cook & Mineka (1987, 1989, 1990) have shown that rhesus monkeys learn a relationship between snake stimuli and conspecific fear more readily than a relationship between flower stimuli and conspecific fear. If this effect is due to the content, or identity, of the stimuli (snakes *vs.* flowers), rather than to the differential salience of the snake and flower stimuli employed (e.g. differences in colour or brightness), then the effect suggests that observational conditioning of snake fear in rhesus monkeys is, at minimum, an atypical example of stimulus-stimulus learning. 'Selectivity associability' of this kind has been reported in the literature on asocial learning, but it is rare (LoLordo, 1979; LoLordo & Droungas, 1989).

The strongest evidence that monkeys' superior learning of a snake-fear relationship is due to the identity, rather than the differential salience, of the snake and flower stimuli, came from an experiment with monkeys in which the S<sub>2</sub> was food, rather than conspecific fear (Cook & Mineka, 1990, Experiment 3). On each trial in this experiment, a monkey was shown one of four pairs of stimuli on a video screen, and if it reached towards one of the stimuli (+) it was rewarded with food, while reaching towards the other was not rewarded (-). Thus, the monkeys had four discrimination problems: (1) coiled snake+/red square-(snake positive); (2) long snake-/red square+(snake negative); (3) chrysanthemums+/blue diamond-(flower positive); (4) silk flowers-/blue diamond+(flower negative).

Seven of the 13 monkeys in the experiment failed to solve all of the problems. That is, they did not develop a reliable tendency to reach towards the stimuli related to reward. However, the remaining monkeys solved the flower problems at least as fast as the snake problems, and in some cases the flower problems were solved more quickly. This was interpreted as evidence that the snake stimuli were no more salient than the flower stimuli, and therefore that superior observational conditioning of snake fear is a

genuine example of selective associability (Cook & Mineka, 1990). However, this conclusion is not secure because (i) the snake and flower stimuli in this experiment were not identical to those used to test observational conditioning of fear, and (ii) the monkeys were required to discriminate snakes and flowers from different, arbitrary stimuli. It may have been more difficult to discriminate the red square from the snake stimuli, than to discriminate the blue diamond from the flower stimuli and, if this was the case, the monkeys might have solved the snake problems more slowly even if the snake stimuli were more salient than the flower stimuli.

(b) *Immunization*

It has also been suggested that observational conditioning of snake fear in rhesus monkeys is unlike other, documented examples of stimulus-stimulus learning in being subject to attenuation through pre-exposure to a conspecific behaving non-fearfully in the presence of snakes (Mineka & Cook, 1986). In the experiment demonstrating this 'immunization' effect, monkeys received one of three treatments before observing a conspecific behaving fearfully in the presence of snakes. The 'immunization' group observed a conspecific reaching non-fearfully across snake stimuli to obtain food. The 'pseudoimmunization' group observed a conspecific reaching non-fearfully for food across neutral stimuli (wood blocks), and members of the 'latent inhibition' group were themselves allowed to reach for food across snake stimuli. When they were tested after observing a demonstrator behaving fearfully in the presence of snakes, the pseudo-immunization group showed more fear of snakes than the other two groups, and there was some evidence that the latent inhibition group were more fearful of snakes than the immunization group.

Mineka & Cook (1986) suggested that this effect is unprecedented in the literature on asocial, stimulus-stimulus learning: "the classical conditioning literature provides no strong empirical grounds from which to predict that watching a model behave nonfearfully will immunize more effectively than will simple exposure alone" (p. 308). This may be so, but the animals in Mineka & Cook's latent inhibition group were not only preexposed to the snake stimuli, but also to a relationship between snake stimuli and food, and this is known within the literature on asocial learning to reduce the retardation of aversive conditioning produced by stimulus preexposure (Dickinson, 1977; Mackintosh, 1983). For example, Dickinson (1976) found that rats which had been preexposed to a tone-food relationship (equivalent to Mineka & Cook's 'latent inhibition' group), subsequently learn a tone-electric shock relationship more slowly than rats that were preexposed to uncorrelated presentations of the tone and food (immunization), but faster than animals that had not been preexposed to the tone (pseudoimmunization).

The experiments discussed here, and others by Cook & Mineka (e.g. 1987), are unique in the refinement with which they have identified and addressed questions about the mechanisms underlying social learning. However, they have not revealed any substantial differences between social and asocial stimulus-stimulus learning. On the contrary, they underline similarities between the conditions of learning in the two cases, and thereby encourage the hypothesis that the same mechanisms are responsible.



*(c) Spatial contiguity*

There is no compelling evidence in the literature on asocial learning in animals that exposure to a spatial relationship between events, in the absence of a temporal relationship, will give rise to learning (Rescorla, 1980). However, studies of observational conditioning in neonatal domestic chicks (Suboski & Bartashunas, 1984) have been taken to show that spatial contiguity can be a sufficient basis for learning (Suboski, 1990). In these studies, day-old chicks observed a motor-driven arrow (a schematic version of an adult hen) 'pecking' one of an array of coloured pinheads. When the birds were, simultaneously or subsequently, presented with similar array of pinheads, they tended to peck those of the same colour as the pinhead pecked by the arrow.

This result was interpreted as evidence that the chicks' behaviour was affected by spatial contiguity between the pecked pinhead (S<sub>1</sub>) and the pecking arrow (S<sub>2</sub>), in the absence of a temporal relationship, because the experimenter did not arrange for the pecked pinhead to have a stronger temporal relationship than the other pinheads with the pecking arrow (Suboski, 1990). However, regardless of whether the experimenter planned it to be so, the temporal relationship between the pecked pinhead and the pecking arrow may have been greater in the subjects' perception. It is possible that the chicks looked at the pecked pinhead just before looking at the arrow more often than they looked at the other pinheads just before looking at the arrow. Consequently, the chicks' preference for stimuli matching the pecked pinhead could have resulted from exposure to a temporal, rather than an exclusively spatial, relationship.

*(d) Response evocation vs. learnability*

Experiments on observational conditioning of food preferences in rats suggest that this phenomenon may be unlike many examples of asocial, stimulus-stimulus learning, in being detectable through changes in response evocation, but not learnability. Galef and his associates (Galef & Wigmore, 1983; Galef, Kennett & Wigmore, 1984; Galef, Kennett & Stein, 1985; Galef & Stein, 1985) found that rats which have been exposed to the odour and/or flavour of a food (S<sub>1</sub>) on the breath of a conspecific (S<sub>2</sub>) subsequently consume more of the S<sub>1</sub> food than of other diets. Thus, they detected an effect of socially mediated exposure to a stimulus-stimulus relationship through an increase in the extent to which the S<sub>1</sub> evoked a response. In the case of asocial learning, effects of stimulus-stimulus exposure are commonly detectable using both a response evocation measure and a learnability measure. Therefore, if social and asocial learning are alike, one would expect the effects of exposure to one diet (S<sub>1</sub>) with conspecific breath (S<sub>2</sub>) to be detectable in the degree to which the observer learns about the relationship between another diet (S<sub>3</sub>) and conspecific breath (S<sub>2</sub>). However, Galef & Durlach (1993) have failed to find overshadowing and blocking effects of this kind. For example, they showed that rats which had experienced marjoram-flavoured diet (S<sub>1</sub>) on a demonstrator's breath (S<sub>2</sub>), and then cinnamon-flavoured diet (S<sub>3</sub>) on a demonstrator's breath (S<sub>2</sub>), subsequently consumed the same amount of cinnamon-flavoured diet as observers that had experienced the S<sub>3</sub>-S<sub>2</sub> pairing, but not the S<sub>1</sub>-S<sub>2</sub> relationship.

As Galef & Durlach (1993) acknowledge, it would be premature to conclude that this observational conditioning phenomenon is unlike other examples of stimulus-stimulus learning. Overshadowing and blocking effects are not invariably found in studies of asocial, stimulus-stimulus learning (Durlach & Rescorla, 1980; Williams, 1981; Dickinson, Nicholas & Mackintosh, 1983), and there is reason to believe that they may occur more readily in multiple-trial procedures (Rescorla & Wagner, 1972; Mackintosh, 1975) than in the kind of single-trial procedure used by Galef & Durlach (1993). However, the possibility that observational conditioning of food preferences in rats is detectable only through response evocation is intriguing, and worthy of further investigation.

#### IV. OBSERVATIONAL LEARNING AS R-S LEARNING

##### (1) *Current definitions: imitation, copying and observational learning*

The term *imitation* has been used in a number of different ways in the literature on social learning. However, it is now commonly understood to refer to the acquisition of a topographically novel response through observation of a demonstrator making that response, and to involve a 'goal-directed' psychological mechanism (Galef, 1988). Thus, in the case of imitation, conspecific observation is understood to be sufficient for acquisition of the novel response, and the demonstrator is said to 'model' the response.

Compelling demonstrations of imitation in animals are rare, but one example was provided by Galef, Manzig & Field (1986) in an experiment with budgerigars. On each of a number of trials, the birds in this study observed a conspecific demonstrator lifting a flat cover from a food dish with its feet or with its beak, and were then allowed access to the cup-and-cover apparatus themselves. Birds that had observed the demonstrator using its beak were more likely to use their beak than their feet to lift the cover, and vice versa for birds that had observed a demonstrator using its feet.

Imitation is sometimes contrasted with *copying*. In this kind of social learning, a novel response is acquired through observation of a demonstrator making that response, but observation of the demonstrator's behaviour is not sufficient for response acquisition, and a 'goal-directed' psychological mechanism is not thought to be involved (Galef, 1988). Instead, it is assumed that the observer exhibits components of the matching behaviour by chance, and that, after observation, these components increase in frequency because they are either selectively rewarded by a human observer (Millar & Dollard, 1941), or intrinsically rewarding for the observer to execute (Thorndike, 1911; Mowrer, 1960). In the former case, it is sufficient for a human observer to be sensitive to the degree of similarity between the observer's and the demonstrator's behaviour, but in the latter case the observer animal must have this sensitivity.

Copying is invariably exemplified with reference to the reproduction by birds of conspecific song and/or human vocalizations. Consequently, copying is virtually synonymous with *vocal imitation*, and the term *motor imitation* is sometimes used as a substitute for *imitation*.

*Observational learning* is regarded by some authors as synonymous with *imitation* (e.g. Galef, 1988), and by others as a generic term, equivalent to *social learning* (e.g. Osgood, 1953; Hall, 1963).

*(2) Problems*

Isolation of what is currently known as imitation (or observational learning) has been the purpose of most attempts to differentiate categories of social learning. The contemporary classification of social learning (summarized by Galef (1988) and reviewed in the present article) emerged in accordance with, and in reaction to, distinctions made by Morgan (1900), Thorndike (1911) and Thorpe (1956) (see Galef, 1988 for an historical view), and their primary aim in making these distinctions was to circumscribe 'reflective', 'deliberate' or 'intentional' (Morgan, 1900, p. 193) social learning; examples of social learning in animals which, they suspected, could not be mediated by the mechanisms responsible for most asocial learning.

Thus, a desire to identify a distinctive mechanism of social learning has motivated the definition of imitation, and the classification of social learning more generally. It is therefore ironic that the product of these enterprises, a definition of imitation as a phenomenon in which topographically matching behaviour is acquired through a goal-directed mechanism, has obstructed rather than facilitated investigation of its underlying mechanisms. The practice of defining imitation in relation to a goal-directed mechanism has obstructed research on the mechanism of imitation in two ways: first, since a mechanism is something that cannot be observed directly, and it is not clear how the operation of a goal-directed mechanism can be detected in behaviour, it has proved extraordinarily difficult to identify specimens of imitation for analysis, i.e. examples of imitation in animals. With rare exceptions (Dawson & Foss, 1965; Galef, Manzig & Field, 1986; Heyes & Dawson, 1990), whenever a putative example of imitation has been put forward, it has subsequently been suggested or demonstrated that the phenomenon involves one of the non goal-directed processes assumed to underlie copying, local enhancement, stimulus enhancement, or one of the other current categories of social learning (e.g. Galef, 1988; Visalberghi & Fragaszy, 1990). Second, the practice of defining imitation in relation to its mechanism has concealed the need for research. It gives the impression that the mechanism of imitation is fully understood, while, in truth, knowledge of this mechanism consists of an untested hypothesis stating that it has an undefined property, namely, goal-directedness (Davis, 1973; Galef, 1988).

*(3) Proposal*

It is proposed that observational learning be regarded as a subset of response-reinforcer learning (R-S) in which observation of a demonstrator exposes the observer to a relationship between a response and a reinforcer at  $t_1$ , and exposure to this relationship effects a change in the observer detected, in any behaviour, at  $t_2$ . Thus, observational learning differs from asocial response-reinforcer learning in that it is the demonstrator, not the learner, who makes the response that is learned. It is further suggested that imitation be regarded as a special case of observational learning in which the effect of exposure to a positive relationship between a demonstrator's response and appetitive reinforcement at  $t_1$  is detected in the production of novel, topographically matching behaviour by the observer at  $t_2$ .

The advantage of the proposed definition of imitation is that it would circumscribe those social learning phenomena taken by previous investigators to be goal-directed,

while allowing those phenomena, specimens of imitation, to be identified with certainty, and subjected to causal analysis without undue prejudice. To demonstrate that a social learning phenomenon is an example of imitation, it would be sufficient to show that exposure to a positive relationship between a demonstrator's response and appetitive reinforcement at  $t_1$ , resulted in the production of novel, topographically matching behaviour by the observer at  $t_2$ . It would not be necessary to show, in addition, that the effect was mediated by a goal-directed process. On the contrary, having identified examples of imitation, researchers would be free to generate and test hypotheses concerning the mechanisms responsible, including, of course, the hypothesis that the mechanism underlying imitation is goal-directed.

Distinguishing observational learning and imitation in the proposed manner raises the possibility that, in addition to imitation, there are a number of observational learning phenomena which have not yet been identified or documented. If, as in the case of imitation, learning can result from socially mediated exposure to a positive relationship between a response and an appetitive stimulus (a reward contingency), then it may also result when the socially mediated exposure is to a positive relationship involving an aversive stimulus (a punishment contingency), a negative relationship involving an appetitive stimulus (an omission contingency), or a negative relationship involving an aversive stimulus (an avoidance contingency). Furthermore, any change that results from socially mediated exposure to one of these contingencies may be detectable, not only through a change in the probability of novel matching behaviour, but also through its effects on novel non-matching behaviour, or on established components of the observer's behavioural repertoire that do or do not match those of the demonstrator. Generally speaking, one might expect socially mediated exposure to reward and avoidance contingencies to increase the probability of matching behaviour, and socially mediated exposure to punishment and omission contingencies to reduce that probability.

The results of a recent experiment lend support to the hypothesis that animals may be able to learn about a variety of response-reinforcer relationships through conspecific observation (Heyes, Jaldow & Dawson, in press). The rats in this experiment were first trained to push a joystick in one of two directions for food reward, and were then allowed to observe a conspecific demonstrator either sitting passively in the joystick chamber (Group NONE) or pushing the joystick without receiving food reward. Half of the rats in the latter group observed the demonstrator pushing the joystick in the direction to which they, the observer, had learned to push (Group SAME), and half observed the demonstrator pushing in the opposite direction (Group DIFFERENT). After observation, the observers were given an extinction test in which they were allowed to push the joystick but their responses were not rewarded. During the test, the observers in Group SAME made fewer previously rewarded responses than those in Group DIFFERENT, and those in Group DIFFERENT made fewer previously rewarded responses than those in Group NONE. Thus, in Group SAME an effect of socially mediated exposure to a negative relationship between a response and an appetitive reinforcer was detected through a change in an established, matching component of the observer's behavioural repertoire.

(4) *Anomalies*

In this section, an attempt will be made to answer three potential objections to the proposed alignment of observational learning and imitation with response-reinforcer learning.

(a) *Matching disparate sensory inputs*

It might be argued that defining imitation as a type of response-reinforcer learning is at odds with an interesting feature of imitation, namely, that it seems to require the observer to equate its own behaviour with that of a demonstrator when the two give rise to disparate sensory stimulation (Heyes, in press). For example, budgerigars apparently equate the movements of a conspecific's beak with movements of their own beak (Galef *et al.*, 1986), and yet the visual and kinaesthetic feedback from a bird's own beak movements must be very different from the sensory stimulation produced by a demonstrator's beak movements.

Little or nothing is known of the mechanisms responsible for the matching of disparate sensory inputs required for imitation, and the requirement itself certainly distinguishes imitation from asocial response-reinforcer learning. However, it is not clear why either of these considerations might render inappropriate the treatment of imitation as a type, or special case, of response-reinforcer learning. Such treatment constitutes an hypothesis that the mechanisms responsible for social and asocial response-reinforcer learning are the same save that certain examples of social response-reinforcer learning, those called imitation, involve *additional* mechanisms equating observed and executed responses. Future experiments may disconfirm this hypothesis and thereby reveal that imitation is not a variety of response-reinforcer learning. But the simple fact that imitation requires matching of disparate sensory inputs does not, in itself, constitute disconfirming evidence.

(b) *Intrinsic reinforcement*

The alignment of observational learning with response-reinforcer learning assumes that socially mediated exposure to a response will not lead to a change in behaviour unless that response is related in the observer's experience with a reinforcer. However, there are examples of observational learning, those conventionally described as examples of copying, in which a change in the observer's behaviour occurs when neither the demonstrator's behaviour nor the observer's matching behaviour is apparently reinforced. For example, parrots apparently reproduce human vocalizations when neither the human's nor the parrot's matching vocalizations are paired with appetitive stimulation.

This assumption or, more accurately, hypothesis, is consistent with existing, 'intrinsic reinforcement' theories of copying or vocal imitation (Thorndike, 1911; Mowrer, 1960). Loosely interpreted, these theories suggest that, while copying does not involve overt or extrinsic reinforcement, chance production of the demonstrator's behaviour leads the observer to generate a stimulus internally, and that exposure to the relationship between the matching behaviour and this stimulus, the intrinsic reinforcer, is necessary to produce a durable change in behaviour. However, these theories have barely been tested and, until they are, the hypothesis that exposure to a

response–reinforcer relationship is necessary for observational learning will remain one of the most contentious and therefore challenging, features of the proposed classification of social learning.

(c) *Matched-dependent behaviour*

Matched-dependent behaviour is conventionally regarded as a distinctive type of social learning, and one that is related to operant conditioning or response–reinforcer learning (Galef, 1988). Consequently, some readers may be surprised that matched-dependent behaviour has not been mentioned in the foregoing discussion of observational learning. Before this omission is explained, the origin of the term matched-dependent behaviour will be given.

Miller & Dollard (1941) conducted experiments in which rats were repeatedly given a choice of two alleys to enter in a T-maze. The rats observed a conspecific entering one of the two alleys just before making their choice, and if they entered the same alley as the demonstrator, they were rewarded. Thus, it is claimed, the demonstrator acted as a 'discriminative stimulus' indicating which of the two responses would be rewarded on that trial. Over trials, the observers acquired a tendency to enter the same alley as their demonstrators, and the term *matched-dependent behaviour* is used to describe this tendency.

Matched-dependent behaviour was omitted from the foregoing discussion because it does not appear to be a form of social learning. (For the same reason, reference has not been made to social facilitation or contagious behaviour.) In the experiments by Miller & Dollard (1941) which define the category by ostension, the effects of demonstrator observation at  $t_1$  were detected at  $t_1$ , not at  $t_2$  when the demonstrator was no longer present. Instead of demonstrating or analysing a social learning phenomenon, Miller & Dollard examined conditions that may promote the occurrence of social learning. Following other animals is likely to increase the probability of social learning, and Miller & Dollard (1941) showed that, in circumstances where rats do not spontaneously follow conspecifics, they will acquire a tendency to do so if following is differentially rewarded. In a similar vein, Hogan (1986, 1988) has shown that pigeons can be trained to attend selectively either to a keylight of the same colour as that pecked by a conspecific, or to a keylight of a different colour.

Evidence that acquired following, or matched-dependent behaviour, can lead to social learning was provided by Church (1957). On each trial in the first stage of his experiment, rats were rewarded if they entered the same arm of a T-maze as a conspecific 'leader', and not rewarded if they entered the alternative arm. In the second stage, the observer rats were again rewarded for following their leaders, but an incidental cue was added: a red light shone in the alley entered by the leader, but not in the other alley. In the final stage, the leaders were absent, but the rats continued to enter the alley marked with a red light. Thus, the rats appear to have learned, as a consequence of matched-dependent behaviour, about the relationship between two stimuli, the red light and food. Unlike matched-dependent behaviour, this is a social learning phenomenon in its own right, but it is an example of observational conditioning, S–S learning, not observational learning, R–S learning.

## V. SUMMARY AND CONCLUSIONS

There has been relatively little research on the psychological mechanisms of social learning. This may be due, in part, to the practice of distinguishing categories of social learning in relation to ill-defined mechanisms (Davis, 1973; Galef, 1988). This practice both makes it difficult to identify empirically examples of different types of social learning, and gives the false impression that the mechanisms responsible for social learning are clearly understood.

It has been proposed that social learning phenomena be subsumed within the categorization scheme currently used by investigators of asocial learning. This scheme distinguishes categories of learning according to observable conditions, namely, the type of experience that gives rise to a change in an animal (single stimulus *vs.* stimulus-stimulus relationship *vs.* response-reinforcer relationship), and the type of behaviour in which this change is detected (response evocation *vs.* learnability) (Rescorla, 1988). Specifically, three alignments have been proposed: (i) stimulus enhancement with single stimulus learning, (ii) observational conditioning with stimulus-stimulus learning, or Pavlovian conditioning, and (iii) observational learning with response-reinforcer learning, or instrumental conditioning.

If, as the proposed alignments suggest, the conditions of social and asocial learning are the same, there is some reason to believe that the mechanisms underlying the two sets of phenomena are also the same. This is so if one makes the relatively uncontroversial assumption that phenomena which occur under similar conditions tend to be controlled by similar mechanisms. However, the proposed alignments are intended to be a set of hypotheses, rather than conclusions, about the mechanisms of social learning; as a basis for further research in which animal learning theory is applied to social learning.

A concerted attempt to apply animal learning theory to social learning, to find out whether the same mechanisms are responsible for social and asocial learning, could lead both to refinements of the general theory, and to a better understanding of the mechanisms of social learning. There are precedents for these positive developments in research applying animal learning theory to food aversion learning (e.g. Domjan, 1983; Rozin & Schull, 1988) and imprinting (e.g. Bolhuis, de Vos & Kruit, 1990; Hollis, ten Cate & Bateson, 1991). Like social learning, these phenomena almost certainly play distinctive roles in the ontogeny of adaptive behaviour, and they are customarily regarded as 'special kinds' of learning (Shettleworth, 1993). However, recent analyses have shown that they occur through associative mechanisms akin to those underlying laboratory demonstrations of Pavlovian conditioning and perceptual learning. Perhaps, this review suggests, it is time also to reexamine the 'special' status of social learning.

## VI. ACKNOWLEDGEMENTS

This research was supported by The Leverhulme Trust. I am grateful to Anthony Dickinson, Bennet Galef, Henry Plotkin, Phil Reed and an anonymous reviewer for their valuable comments on an earlier draft of the manuscript.

## VII. REFERENCES

- BEAUCHAMP, G. & KACELNIK, A. (1991). Effects of the knowledge of partners on learning rates in zebra finches. *Animal Behaviour* **41**, 247-253.

- BOLHUIS, J. J., DE VOS, G. J. & KRUIJT, J. P. (1990). Filial imprinting and associative learning. *Quarterly Journal of Experimental Psychology* **42B**, 313-329.
- BOX, H. O. (1984). *Primate Behavior and Social Ecology*. Chapman & Hall, London.
- CHURCH, R. M. (1957). Transmission of learned behavior between rats. *Journal of Abnormal and Social Psychology* **54**, 163-165.
- COOK, M. & MINEKA, S. (1987). Second-order conditioning and overshadowing in the observational conditioning of fear in monkeys. *Behavior Research and Therapy* **25**, 349-364.
- COOK, M. & MINEKA, S. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *Journal of Abnormal Psychology* **98**, 448-459.
- COOK, M. & MINEKA, S. (1990). Selective associations in the observational conditioning of fear in rhesus monkeys. *Journal of Experimental Psychology* **16**, 372-389.
- COOK, M., MINEKA, S., WOLKENSTEIN, B. & LAITSCH, K. (1985). Observational conditioning of snake fear in unrelated rhesus monkeys. *Journal of Abnormal Psychology* **93**, 355-372.
- CURIO, E. (1988). Cultural transmission of enemy recognition by birds. In *Social Learning: Psychological and Biological Perspectives* (ed. T. R. Zentall and B. G. Galef), pp. 75-98. Erlbaum, Hillsdale, NJ.
- DARBY, C. L. & RIOPELLE, A. J. (1959). Observational learning in the rhesus monkey. *Journal of Comparative & Physiological Psychology* **52**, 94-98.
- DAVIS, J. M. (1973). Imitation: a review and critique. In *Perspectives in Ethology* (ed. P. P. G. Bateson and P. H. Klopfer). Plenum, New York.
- DAWSON, B. V. & FOSS, B. M. (1965). Observational learning in budgerigars. *Animal Behaviour* **13**, 470-474.
- DICKINSON, A. (1976). Appetitive-aversion interactions: facilitation of aversive conditioning by prior appetitive training in the rat. *Animal Learning and Behavior* **4**, 416-420.
- DICKINSON, A. (1980). *Contemporary Animal Learning Theory*. Cambridge University Press, Cambridge.
- DICKINSON, A., NICHOLAS, D. J. & MACKINTOSH, N. J. (1983). A re-examination of one-trial blocking in conditioning suppression. *Quarterly Journal of Experimental Psychology* **35B**, 67-79.
- DOMJAN, M. (1983). Biological constraints on instrumental and classical conditioning: implications for general process theory. *Psychology of Learning and Motivation* **17**, 216-277.
- DOMJAN, M. (1993). *The Principles of Learning and Behavior*. Brooks Cole, California.
- DURLACH, P. J. & RESCORLA, R. A. (1980). Potentiation rather than overshadowing in flavor aversion learning: an analysis in terms of within-compound associations. *Journal of Experimental Psychology: Animal Behavior Processes* **6**, 175-187.
- FIORITO, G. & SCOTTO, P. (1992). Observational learning in *Octopus vulgaris*. *Science* **256**, 545-547.
- GALEF, B. G. (1976). Social transmission of acquired behavior: a discussion of tradition and social learning in vertebrates. In *Advances in the Study of Behavior*, vol. 6 (ed. J. S. Rosenblatt, R. A. Hinde, E. Shaw and C. Beer), pp. 77-100. Academic Press, New York.
- GALEF, B. G. (1988). Imitation in animals: history, definition and interpretation of data from the psychological laboratory. In *Social Learning: Psychological and Biological Perspectives* (ed. T. R. Zentall and B. G. Galef), pp. 3-28. Erlbaum, Hillsdale, NJ.
- GALEF, B. G. & BECK, M. (1985). Aversive and attractive marking of toxic and safe foods by Norway rats. *Behavioral and Neural Biology* **43**, 298-310.
- GALEF, B. G. & DURLACH, P. J. (1993). Absence of blocking, overshadowing, and latent inhibition in social enhancement of food preferences. *Animal Learning and Behavior* **21**, 214-220.
- GALEF, B. G., KENNETT, D. J. & STEIN, M. (1985). Demonstrator influence on observer diet preference: effects of simple exposure and the presence of a demonstrator. *Animal Learning and Behavior* **13**, 25-30.
- GALEF, B. G., KENNETT, D. J. & WIGMORE, S. W. (1984). Transfer of information concerning distant foods in rats: a robust phenomenon. *Animal Learning and Behavior* **12**, 292-296.
- GALEF, B. G., MANZIG, L. A. & FIELD, R. M. (1986). Imitation learning in budgerigars: Dawson and Foss (1965) revisited. *Behavioral Processes* **13**, 191-202.
- GALEF, B. G. & STEIN, M. (1985). Demonstrator influence on observer diet preference: analyses of critical social interaction and olfactory signals. *Animal Learning & Behavior* **13**, 31-38.
- GALEF, B. G. & WIGMORE, S. W. (1983). Transfer of information concerning distant foods: a laboratory investigation of the 'information center' hypothesis. *Animal Behaviour* **31**, 748-758.
- HALL, K. R. L. (1963). Observational learning in monkeys and apes. *British Journal of Psychology* **54**, 201-226.
- HEYES, C. M. (in press). Imitation, culture and cognition. *Animal Behaviour* **46**, 999-1010.
- HEYES, C. M. & DAWSON, G. R. (1990). A demonstration of observational learning using a bidirectional control. *Quarterly Journal of Experimental Psychology* **42b**, 59-71.
- HEYES, C. M., JALDOW, E. & DAWSON, G. R. (in press). Observational extinction: observation of non-reinforced responding reduces resistance to extinction in rats. *Animal Learning and Behavior* **21**, 221-225.



- HOGAN, D. E. (1986). Observational learning of a conditional hue discrimination in pigeons. *Learning and Motivation* **17**, 40–58.
- HOGAN, D. E. (1988). Learned imitation by pigeons. In *Social Learning: Psychological and Biological Perspectives* (ed. T. R. Zentall and B. G. Galef), pp. 225–238. Erlbaum, Hillsdale, NJ.
- HOLLIS, K., TEN CATE, C. & BATESON, P. P. G. (1991). Stimulus representation: a subprocess of imprinting and conditioning. *Journal of Comparative Psychology* **105**, 307–317.
- LOLORDO, V. (1979). Selective associations. In *Mechanisms of Learning and Motivation* (ed. A. Dickinson and R. Boakes), pp. 367–398. Erlbaum, Hillsdale, NJ.
- LOLORDO, V. & DROUNGAS, A. (1989). Selective associations and adaptive specializations: food aversion and phobias. In *Contemporary Learning Theories: Instrumental Conditioning Theory and the Impact of Biological Constraints on Learning* (ed. S. Klein and R. Mowrer), pp. 145–179. Erlbaum, Hillsdale, NJ.
- LOLORDO, V. & ROSS, R. T. (1990). Retraction of Ross and LoLordo's findings concerning blocking in serial feature-positive discriminations. *Journal of Experimental Psychology: Animal Behavior Processes* **16**, 402–406.
- LORE, R., BLANC, A. & SUEDFELD, R. (1971). Empathetic learning of a passive avoidance response in domesticated *Rattus norvegicus*. *Animal Behaviour* **19**, 112–114.
- 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*, pp. 100–112. Oxford University Press, Oxford.
- MAYR, E. (1961). Cause and effect in biology. *Science* **134**, 1501–1506.
- MILLER, N. E. & DOLLARD, J. (1941). *Social Learning and Imitation*. Yale University Press, New Haven.
- MINEKA, S. & COOK, M. (1986). Immunization against the observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology* **95**, 307–318.
- MINEKA, S. & COOK, M. (1988). Social learning and the acquisition of snake fear in monkeys. In *Social Learning: Psychological and Biological Perspectives* (ed. T. R. Zentall and B. G. Galef), pp. 51–74. Erlbaum, Hillsdale, NJ.
- MINEKA, S., DAVIDSON, M., COOK, M. & KEIR, R. (1984). Observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology* **93**, 355–372.
- MORGAN, C. LLOYD (1900). *Animal Behaviour*. Edward Arnold, London.
- MOWRER, O. H. (1960). *Learning Theory and the Symbolic Processes*. Wiley, New York.
- OSGOOD, C. E. (1953). *Method and Theory in Experimental Psychology*. Oxford University Press, New York.
- PALAMETA, B. & LEFEBVRE, L. (1985). The social transmission of a food-finding technique in pigeons: What is learned? *Animal Behaviour* **33**, 892–896.
- RESCORLA, R. A. (1980). *Pavlovian Second-Order Conditioning: Studies in Associative Learning*. Erlbaum, Hillsdale, NJ.
- RESCORLA, R. A. (1988). Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience* **11**, 329–352.
- RESCORLA, R. A. & WAGNER, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning, II* (ed. A. Black and W. Prokasy), pp. 64–99. Appleton-Century-Crofts, New York.
- ROZIN, P. & SCHULL, J. (1988). The adaptive–evolutionary point of view in experimental psychology. In *Stevens' Handbook of Experimental Psychology*, 2nd ed., vol. 1 (ed. R. C. Atkinson, R. J. Herrnstein, G. Lindzey and R. D. Luce), pp. 503–546. Wiley, New York.
- SHETTLEWORTH, S. J. (1993). Varieties of learning and memory. *Journal of Experimental Psychology: Animal Behavior Processes* **19**, 5–14.
- SPENCE, K. W. (1937). Experimental studies of learning and higher mental processes in infra-human primates. *Psychological Bulletin* **34**, 806–850.
- SUBOSKI, M. D. (1990). Releaser-induced recognition learning. *Psychological Review* **97**, 271–284.
- SUBOSKI, M. D. & BARTASHUNAS, C. (1984). Mechanisms for the social transmission of pecking preferences in neonatal chicks. *Journal of Experimental Psychology: Animal Behavior Processes* **10**, 182–194.
- THORNDIKE, E. L. (1911). *Animal Intelligence*. Macmillan, New York.
- THORPE, W. H. (1956). *Learning and Instinct in Animals*. Methuen, London.
- VISALBERGHI, E. & FRAGASY, D. M. (1990). Do monkeys ape? In *Language and Intelligence in Monkeys and Apes* (ed. S. Parker and K. Gibson), pp. 247–273. Cambridge University Press, New York.
- WARDEN, C. J. & JACKSON, T. A. (1935). Imitative behaviour in the rhesus monkey. *Journal of Genetic Psychology* **46**, 103–125.
- WHITEN, A. & HAM, R. (1992). On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. In *Advances in the Study of Behavior*, vol. 21 (ed. P. J. B. Slater, J. S. Rosenblatt, C. Beer and M. Milinski). Academic Press, New York.
- WILLIAMS, B. A. (1981). Blocking in an autoshaping procedure. *Behavior Analysis Letters* **1**, 345–351.
- ZENTALL, T. R. & LEVINE, J. M. (1972). Observational learning and social facilitation in the rat. *Science* **120**–1221.