

**Please note**

The text in this file has been automatically extracted and may contain minor errors. For the original version please consult the paper copy held in the Swinburne Library.

# 3

## ACE: BASIC CHARACTERISTICS

OVERVIEW	3-1
HIGHER-ORDER CONDITIONING	3-2
TYPE OF ASSOCIATION SUPPORTED	3-3
CS INPUT CHARACTERISTICS	3-5
US INPUT CHARACTERISTICS	3-6
DRIVE, PERFORMANCE, AND LEARNING	3-8
SUMMARY	3-9

### OVERVIEW

In this **chapter**, the type of association to be supported by ACE, the nature of its **inputs**, and the nature of its **output** are established. The basic characteristics of ACE differ markedly from those of previous **neuronal** models of classical conditioning. Issues concerning the internal organisation of ACE are covered in later chapters.

Unlike most other neuronal models of classical conditioning, ACE **forms** an association between the CS and a representation of the US, rather than between the CS and the CR. Consequently, the output from ACE is not equivalent to **the** CR, **but** instead is best regarded as an indicator of the expectation of the US by the *CS*. This output signal is suitable for subsequent generation of the CR, and as a source of conditioned reinforcement for both higher-order conditioning and instrumental conditioning. Considerable empirical evidence in support of this concept is presented.

**HIGHER-ORDER CONDITIONING**

It **has** become routine for neuronal models of classical conditioning to directly support both first-order **and** higher-order conditioning. This is usually achieved by **implementing** a literal interpretation of stimulus equivalence in which a CS **and** a US are indistinguishable **as** sources of reinforcement and response production. Both **Klopf (1987) and Barto and Sutton (1985)** attribute the US input with the same effect upon the output of their models as an excitatory CS, and the consequent changes in output **signal as** the sole source of reinforcement.

Aside from **minimising** system complexity, the main advantage of this **approach** is that a simple form of higher-order conditioning is automatically supported by whatever **learning** rules **are** developed to support primary conditioning. However, significant differences exist between the empirical characteristics of higher-order and first-order conditioning, which suggest that such **total** stimulus and reinforcer equivalence is ill conceived.

**Gormezano (1984)** notes that a US is considerably **more** effective than an excitatory CS as a **source** of reinforcement at longer ISIs. **Holland and Rescorla (1975a, b)** found that the performance of a higher-order CS can be **essentially** independent of drive level, while **the** performance of a first-order CS appears to be strongly modulated by drive. In addition, the type of CR able to be easily conditioned to a **CS** is often quite restricted for first-order conditioning, but is much less restricted in higher-order conditioning (**Rescorla, 1984**). **Also**, "the nature of the CR elicited by a CS paired with a US does not **seem** to have **any** effect on the nature of the CR it is **able** to establish to a second CS as a result of higher-order conditioning." (**Mackintosh, 1983, p. 16**).

**These** observations suggest that the locus of higher-order association is functionally distinct from that of first-order conditioning, effectively precluding its inclusion within a single neuronal model of first-order classical **conditioning**.

## TYPE OF ASSOCIATION SUPPORTED

Mackintosh (1983) argues that "A theory of [classical] conditioning must first of all be a theory of how associations are established between stimuli ... and reinforcers, and only then a theory of how associations are translated into performance. The two questions are distinct, and attempts to ignore the distinction have provided **bad** answers to both questions." (p. 19). The reasoning used to **reach** this conclusion can be summarised as follows:

Formation of simple **S-R** (Stimulus-Response) associations alone should be rejected because they cannot account for the ability of a **CS** to reinforce new conditioning (either higher-order or instrumental) in terms of its corresponding **CR**. This is primarily because the nature of the higher-order or instrumental response subsequently conditioned can be very different to the **CR** elicited by the **CS**. Mackintosh cites several experimental results in support of this (Holland, 1977; Leyland, 1977; Nairne and Rescorla, 1981). The implication is that the **reinforcing** ability of a **CS** cannot be dependent upon its **CR**, but instead may depend upon an association between the **CS** and the original **US**. **This CS-US** association would then be **followed** by a subsequent association with a particular **CR**, in order to translate the **CS-US** association into performance of the **CR**. Even though a single ACE does not directly support either higher-order or instrumental conditioning, it does need to support classical conditioning in a manner which facilitates this behavior when interacting with other elements within a network. For example, "A first-order **CS** paired with food can be used not only to establish second-order classical conditioning to a second **CS**, but **also** to serve as a conditioned reinforcer for instrumental responses on whose occurrence it is made contingent" (Mackintosh, 1974, p. 89).

Further support for a **CS-US** rather than a **CS-CR** association comes from the results of specially devised blocking experiments (Blanchard and Honig, 1976; Leyland and Mackintosh, 1978; Holland, 1977). Mackintosh (1983, p. 18) notes how these indicate that the ability of one **CS** to **block** conditioning to a second **CS** depends upon the use of a common reinforcer (**US**) for each, and not on **the** similarity of their **CRs**, or their similarity to **CRs** potentially conditionable

to the second CS. Since ACE is to directly support blocking phenomena, this also strongly suggests that ACE should form CS-US associations.

It would therefore be a mistake to make the output of a conditioning element such as ACE correspond to either the UR, or the CR. Indeed, there is considerable evidence against the formation of direct CS-UR associations (Mackintosh, 1983, pp. 51-56). Instead, ACE's output **signal** should be regarded as a "CS expects US" output, which is subsequently converted into a CR by further ANN elements.

Note that this conclusion is very different to that implicitly assumed in the neuronal models of both Klopff (1987) and Barto and Sutton (1985), which regard the output from their elements as corresponding to both the UR and the CR. This seems a most unlikely solution, since Soltysik (1971) found that increasing satiation **may** have a drastic effect **upon** a salivary CR, but little affect upon the UR - which is difficult to achieve when both are united at the point at which the CR is first **adaptively** derived. Furthermore, as discussed above, they seem unaware of the mounting indirect evidence which suggests that reinforcer equivalence is better accounted for by inter-element interactions, rather than intra-element processes. Perhaps the lure of a potentially simple solution which directly produces both responses, and which also facilitates the ability of each to act as a common source of reinforcement, diverted the attention of Klopff (1987) and Barto and Sutton (1985) away from these complexities.

If ACE's output does not correspond with the UR, then additional ANN elements are **required** to produce the UR in response to the US. It follows that, unlike the neuronal models of both Klopff (1987) and Barto and Sutton (1985), the US input need not produce an output response from ACE. Indeed, a more flexible relationship between the CR and the UR can be facilitated by **not** allowing the US to directly activate ACE.

If the US input does not activate ACE's output, then changes in output activation alone no longer contain sufficient information to support a

mechanism of reinforcement of CS-US associations. Thus, a new reinforcement mechanism **also** needs to be developed,

In summary, ACE is to support the formation and **maintenance** of associations between conditioned stimuli and reinforcement, with its output providing a measure of the positive expectation of reinforcement. Subsequent ANN **elements**, to be developed by research other than that described herein, will be responsible for converting this expectation of reinforcement into the performance of a specific CR. The US input to ACE will function only as a source of reinforcement, and **will not be able** to directly **energise ACE's** output. **All of these fundamental** characteristics differentiate ACE from the **neuronal** models of **Klopf (1987) and Barto and Sutton (1985)**.

### **CS INPUT CHARACTERISTICS**

From the experimenter's point of **view**, both the CS **and** the US are conceptually discrete sensory events. However, the subject is presented with **numerous component** stimuli, from both **the** explicitly scheduled CS and US events, and the **experimental** context. From **this** jumble of stimuli, the subject needs **to** selectively **associate** particular **stimulus** components with some aspect **of** the US (**discussed** further below) **based upon** the contingencies it experiences, its prior experience, and any **phylogenetically** determined preferences **for associating** particular types of events.

An ANN element, operating within a distributed **system**, is effectively precluded from being presented with a neatly predefined CS input by the potential diversity of a CS. Instead, an ANN **element** is more likely to be presented with an array of different component CS inputs, from which it may selectively associate with the US representation of the element, those which most reliably predict an unexpected subsequent change in US availability. A unified CS representation **will** usually then be formed during acquisition of the CS-US association. A subject given sufficient opportunity to discriminate a CS **from** other stimuli, **will** associate those stimulus aspects which distinguish the CS **from** other **stimuli**, with the subsequent unexpected changes of reinforcement.

The above considerations suggest that many potential CS stimulus features should converge upon a single ACE, so that some of those features which differentiate the CS from other stimuli **may** come to be associated with the US. However, a single **ACE** need not become dedicated to a single CS, for it would then be unable to support the blocking phenomenon observed amongst multiple **CSs** which share a common reinforcer. Hence, the multiple **CS** stimulus features will activate a mixture of different, possibly incomplete and overlapping, learnt CS definitions, each consisting of at least one **CS** input. Each **CS** stimulus feature corresponds to a "CS input" to ACE. For the purposes of testing, a single CS input is often sufficient to demonstrate **many** types of conditioning behavior, and so **can** be considered to correspond to a complete external CS event. **However, in** practice, and for the demonstration of some specific types of behavior (such as overshadowing and blocking), multiple CS inputs will be simultaneously activated.

#### US INPUT CHARACTERISTICS

Unlike a CS, a US requires no explicit experimental conditioning in order to produce its UR. Indeed, this is **why** Pavlov referred to it as the Unconditioned Stimulus, and its response as the Unconditioned Response (Pavlov, 1927, p. 25). It is therefore reasonable to assume that **unlike** a CS, an **internal** classification of the US has already been **formed**. It remains to consider how specific this classification is to a particular US. Two **main possibilities** exist, corresponding to two types of **classical** conditioning, which **are** distinguished by the type of the UR they produce (**Konorski**, 1967). The first, termed "preparatory conditioning" produces URs which are diffuse expressions of a **general** emotional state, such as approaching an **appetitive** reinforcer, **withdrawing** from an aversive one, excitement, or suppression of ongoing activity. Preparatory **URs** are relatively insensitive to the particular stimulus attributes of the US. For example, a shock administered in many different ways will produce a **similar** preparatory response.

The second type of classical conditioning is "consummatory conditioning", which produces well defined discrete reflexive responses, such as salivation, blinking and pecking. In contrast to preparatory conditioning, consummatory conditioning produces responses which **are** very dependent upon the precise nature of the US. For example, a shock delivered to the cheek will produce an eye-blink, while one delivered to the paw will produce a leg flexion UR.

Konorski (1967) **assumed** that in most **cases** both preparatory and consummatory conditioning **occur**. Reasonably consistent **empirical** evidence suggests that **this** distinction is indeed valid, that preparatory conditioning is a **necessary** precursor to **consummatory** conditioning, and that **preparatory** conditioning involves an association between a CS **and** a **centralised** emotional representation of the type of US, while consummatory conditioning involves association between a CS and precise sensory attributes **of** the US (Mackintosh, 1983, pp. 56-62).

It **was** necessary to commit ACE to one type of conditioning, to enable a closer match with specific empirical **results**, and **to limit** the **scope** of research. In the case of classical conditioning, the nature of the CR will **usually** be similar to that of the UR, **and** so will also be of **similar** specificity. Consummatory conditioning **was** selected, because its very nature produces more specific results which facilitate system design and evaluation. The single US input to ACE **can** therefore be considered to be a highly specific preformed classification of a US, and it is this with which newly formed CS classifications **become** associated.

**DRIVE, PERFORMANCE, AND LEARNING**

There is an apparent tendency for drive level to modulate the performance of classically conditioned responses. For example, a satiated subject produces a weaker salivary CR than an **unsatiated** subject (Soltysik, 1971). However, it is unclear if drive directly modulates CR performance, or if instead it exerts some indirect influence.

The effect of drive level upon learning (**as** opposed to **performance**) is even less clear. It appears that drive level exercises an indirect influence upon learning via its direct **modulation** of performance (Kimble, 1961, p. 413). In the absence of any strong contradictory evidence, it will be assumed **that** drive level does not directly modulate learning. Intuitively, this has the effect of enabling the acquisition of associations when drive level is low, which will then be **available** for use when drive level is high.

It remains to consider if in modulating **performance** of CRs, whether drive level should **modulate** the "CS expects US" output **from** ACE. Clearly, it is not necessary that drive act directly upon ACE's output, as many alternative sites downstream of ACE will be available. Furthermore, if ACE's output was directly modulated by drive level, then so also would be the conditioned reinforcement it supports for both higher-order and instrumental conditioning. Since the author is unaware of any evidence in support of this, it is assumed that drive level does not act directly upon ACE, and therefore that ACE does not require a drive input.

## SUMMARY

ACE is to directly support first-order consummatory classical conditioning. The basic configuration of ACE, in terms of its inputs and its output, is illustrated below.

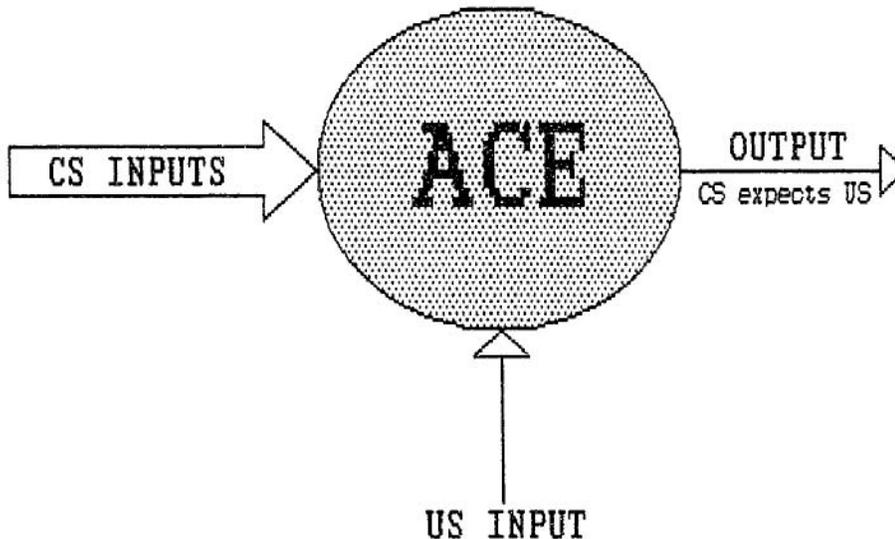


FIGURE 3-1. Basic input/output configuration of ACE.

Multiple CS inputs converge upon ACE where they may become associated with subsequent unexpected changes in US input activity. Each individual CS input corresponds to a stimulus feature which may potentially form part, or all, of a learned CS definition, as a result of conditioning. The US input is highly specific to an individual US event, which is considered to produce a similarly specific UR. Both the US and the UR are assumed to be predefined.

ACE's output does not correspond directly to either the UR or the CR, but instead is best thought of as indicating the strength of the US which is expected by previously presented CSs. This output signal is suitable for subsequent generation of the CR, and as a source of conditioned reinforcement for both higher-order conditioning and instrumental conditioning.