



Bonardi, Charlotte and Robinson, Jasper and Jennings, Dómnall (2017) Can existing associative principles explain occasion setting? Some old ideas and some new data. *Behavioural Processes*, 137 . pp. 5-18. ISSN 1872-8308

**Access from the University of Nottingham repository:**

<http://eprints.nottingham.ac.uk/34998/1/Bonardi%20Robinson%20%20Jennings%202016.pdf>

**Copyright and reuse:**

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution Non-commercial No Derivatives licence and may be reused according to the conditions of the licence. For more details see: <http://creativecommons.org/licenses/by-nc-nd/2.5/>

**A note on versions:**

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact [eprints@nottingham.ac.uk](mailto:eprints@nottingham.ac.uk)

Can existing associative principles explain occasion setting? Some old ideas and some  
new data

Charlotte Bonardi<sup>1</sup>, Jasper Robinson<sup>1</sup> & Dómhnall Jennings<sup>2</sup>

<sup>1</sup> School of Psychology, University of Nottingham

<sup>2</sup> Institute of Neuroscience, Newcastle University

Since occasion setting was identified as a type of learning independent of 'simple' associative processes, a great deal of research has explored how occasion setters are established and operate. Initial theories suggested that they exert *hierarchical* control over a target CS→US association, facilitating the ease with which a CS can activate the US representation and elicit the CR. Later approaches proposed that occasion setting arises from an association between a *configural* cue, formed from the conjunction of the occasion setter and CS, and the US. The former solution requires the associative principles dictating how stimuli interact to be modified, while the latter does not. The history of this theoretical distinction, and evidence relating to it, will be briefly reviewed and some novel data presented. In summary, although the contribution of configural processes to learning phenomena is not in doubt, configural theories must make many assumptions to accommodate the existing data, and there are certain classes of evidence that they are logically unable to explain. Our contention is therefore that some kind of hierarchical process is required to explain occasion-setting effects.

Keywords: occasion setting; hierarchical theory; configural theory; learning.

## Introduction

Skinner (1938) was the first to suggest that performance of an operant response could come under the control of a *discriminative stimulus*. The idea spread to the cognitive behaviourist literature, with Holland (e.g. 1983; 1985) and Rescorla (e.g. 1985; 1986) demonstrating the same effect in Pavlovian conditioning. Specifically, after training that a conditioned stimulus (CS) signalled an unconditioned stimulus (US), animals could confine performance of their conditioned response (CR) to the presence of a stimulus termed a *modulator* (Rescorla, 1985), a *remote initiating stimulus* (Jenkins, 1985) or an *occasion setter* (OS; Holland, 1983).

The truly novel finding was that the control exerted by the OS was *independent of its associative properties*. An OS could enhance performance of a CR that it could not itself elicit (Ross & Holland, 1981; Rescorla, 1985), even after OS extinction (Holland 1989a), and pretraining the OS to signal the US impeded its ability to acquire control over the CR (e.g. Rescorla, 1986)<sup>1</sup>. This suggested that an OS's effect on behaviour did not rely on activation of the US, CS or CR - that it was not mediated by standard associative principles<sup>2</sup>. Two classes of theory emerged in response to this challenge to associative theory, which had until this point dominated accounts of learned behaviour. The first assumed additional, nonassociative principles must be invoked, giving rise to the *US modulation*, *memory systems* and *hierarchical* accounts. The second asserted that existing associative principles could explain occasion setting - provided *combinations* of stimuli could be represented and be subject to associative learning; this class includes the various versions of *configural* theory (e.g. Rescorla, 1972; Brandon, Vogel & Wagner, 2000; Pearce, 1987; 1994). As independent evidence for such configural

---

<sup>1</sup> probably because such pretraining blocked acquisition of associative strength by the target CS (see Swartzentruber, 1995 for a review of related findings)

<sup>2</sup> This should not be taken to imply that an OS may not also have associative properties that influence behaviour, simply that its action cannot be explained solely in those terms.

theories accumulated (e.g., Haselgrove et al., 2008; Pearce, et al. 2002; Williams, Sagness & McPhee 1994), this seemed the more parsimonious explanation, with the result that alternative accounts of occasion setting were eclipsed. This article will revisit evidence relating to these issues, and evaluate the extent to which associative theory can explain occasion setting, or whether additional nonassociative principles are required.

We begin with the *US Modulation and Hierarchical Accounts* theories of occasion setting, for which the evidence, predominantly favouring the hierarchical account, is described. The *Configural Learning* alternative to the hierarchical approach, and its failure to account for evidence of US and CS/S specificity, is then considered; after this the evidence on *Mechanisms of Occasion-Setter Formation* which challenged the hierarchical approach is reviewed. *Elaborations of Hierarchical and Configural Theories* are then presented that can, with added assumptions, explain most of the existing data. Some *Further Discriminating Evidence* that could allow us to choose between these elaborated theories is discussed. We conclude by considering whether or not there is a need to suppose hierarchical processes to explain occasion-setting effects.

### **US Modulation and Hierarchical Accounts**

The first key theories were the US modulation (Rescorla, 1985) and hierarchical accounts (Holland, 1983). *US modulation* elaborated on the existing conceptualisation of a conditioned inhibitor (a stimulus predicting the omission of an otherwise expected US, and counteracting the effect of CSs predicting that US; e.g. Konorski, 1948; Rescorla, 1969) as acting through suppression of activation in the US representation. The US modulation account proposed the complementary process, that a positive occasion setter lowers the activation threshold of the US representation, increasing its sensitivity to excitatory cues. This allows the CS to activate the US representation with greater ease in

the presence of the OS than in its absence. In contrast, the *hierarchical account* asserted that the OS facilitates operation of the association between CS and US (Figure 1<sup>3</sup>).

*Figure 1 about here*

The accounts may be discriminated in terms of *transfer*. Suppose an OS signals that  $cs_1$  predicts  $US_1$  (Figure 1). If the OS facilitates activation of  $US_1$ , it will enhance responding to  $cs_2$  associated with  $US_1$ , but be without effect on  $cs_3$  associated with  $US_2$  - it is *US-specific*, not *CS-specific*. But if the OS enhances operation of the  $cs_1 \rightarrow US_1$  association it will have no effect on  $cs_2$ , even if  $cs_2$  is also associated with  $US_1$ : it will be *both CS-specific and US-specific*. Evidence suggests that occasion setters are both CS- and US-specific, **and that the extent to which such specificity is observed is influenced by procedural factors** (Swartzentruber, 1995).

**CS-specificity** Many studies have shown that  $OS_1$  signalling that  $cs_1$  predicts  $US_1$  ( $OS_1: cs_1 \rightarrow US_1$ ), may control the CR to a  $cs_2$  that *also* predicts  $US_1$ . But this transfer is typically incomplete:  $OS_1$  is rarely as effective with  $cs_2$  as with  $cs_1$  (e.g. Davidson & Rescorla, 1986; Holland, 1986, 1989b, 1989c; Rescorla, 1985; see Swartzentruber, 1995, for a review). This is inconsistent with US modulation: if  $OS_1$  facilitates activation of  $US_1$  it should modulate all CSs associated with  $US_1$  equally. But if  $cs_2$  were to suffer generalisation decrement through being combined with  $OS_1$ , this could reduce responding to  $cs_2$ , allowing US modulation to explain the incomplete transfer. This suggestion has not survived experimental test, however: Bonardi (1996) trained pigeons

---

<sup>3</sup> An alternative conceptualisation of the hierarchical account is that as an occasion-set CS is typically both reinforced and nonreinforced during training, it must have both excitatory and inhibitory associations with the US, and that the occasion setter inhibits the inhibitory association (Bouton & Nelson, 1998). However this account assumes that occasion setting is impossible if the occasion-set CS has no inhibitory strength, and there is evidence against this position (e.g. de Brugada et al., 1995; Hall & Honey, 1989).

that two occasion setters,  $OS_1$  and  $OS_2$  (a tone and flashing houselight) signalled reinforcement of keylights  $cs_1$  and  $cs_2$ , respectively;  $cs_1$  and  $cs_2$  were nonreinforced when presented alone. In contrast  $cs_3$  and  $cs_4$  were reinforced regardless of whether they were accompanied by  $OS_1$  and  $OS_2$  or not (Table 1). Thus  $OS_1$  and  $OS_2$  were occasion setters for  $cs_1$  and  $cs_2$ , but *not* for  $cs_3$  and  $cs_4$ . Then responding to  $cs_1$ ,  $cs_2$ ,  $cs_3$  and  $cs_4$  was examined in the presence of  $OS_1$  and  $OS_2$ , in combinations that were the *same* or *different* from those of training. Incomplete transfer of occasion setting - more responding on *same* ( $OS_1:cs_1$ ,  $OS_2:cs_2$ ) than on *different* ( $OS_1:cs_2$ ,  $OS_2:cs_1$ ) trials - was observed. If this were due to generalisation decrement of  $cs_1$  when it was first presented with  $OS_2$  (and of  $cs_2$  with  $OS_1$ ) then the same effect would be expected on trials with  $cs_3$  and  $cs_4$  - more responding on same ( $OS_1:cs_3$ ,  $OS_2:cs_4$ ) than different ( $OS_1:cs_4$ ,  $OS_2:cs_3$ ) trials. In fact numerically the opposite was observed<sup>4</sup> (Figure 2). If incomplete transfer of occasion setting is not due to generalisation decrement, it implies that occasion setters are CS-specific, contrary to the US-modulation account (cf. Rescorla, 1991a; 1991b).

*Table 1 about here*

*Figure 2 about here*

**Memory Systems** This line of work raised further issues: Holland (1989c) reported that  $OS_1$  signalling that  $cs_1$  predicts  $US_1$  ( $OS_1: cs_1 \rightarrow US_1$ ) modulated responding to  $cs_2$  *only* if it had been trained in an occasion-setting task. Conversely,  $OS_1$  modulated responding to a stimulus that signalled  $US_2$ , but only if  $US_2$  had been involved in an occasion-setting task. Holland (e.g. 1989c) thus proposed the *multiple memory systems* account: (i) events involved in occasion-setting are represented in a different memory system from

---

<sup>4</sup> This may be explained in terms of SOP (Wagner, 1981) as retrieval-generated priming of  $cs_3$  and  $cs_4$  on same trials reducing their ability to elicit the CR - cf., Honey, Hall & Bonardi, 1993.

those that have not, (ii) transfer is more likely between events *within* a memory system. However, many have reported substantial or complete transfer to CSs that have *not* been targets of occasion-setting, creating a problem for this theory (cf., Swartzentruber, 1995).

**Role of Stimulus Generalisation** The literature on transfer reveals great variability in whether transfer of an occasion setter to a different CS is obtained, which is only partly attributable to the transfer CS's training history. Although transfer is often better to CSs that have also been targets of occasion setting, this may be understood in terms of simpler principles such as *stimulus generalisation*. For example, hierarchical theory assumes  $OS_1$ , signalling  $cs_1$  predicts  $US_1$ , acts directly on the  $cs_1 \rightarrow US_1$  link, and so should have no effect on  $cs_2 \rightarrow US_1$ . But if  $cs_1$  and  $cs_2$  comprise unique and common elements (i.e.  $cs_1 = cs_a cs_c$  and  $cs_2 = cs_b cs_c$ ), then both  $cs_1 \rightarrow US_1$  and  $cs_2 \rightarrow US_1$  share a  $cs_c \rightarrow US_1$  component, and  $OS_1$  could influence responding to  $cs_2$  via its effect on this  $cs_c \rightarrow US_1$  association. That transfer is found more often in pigeons with visual keylight CSs than in rats with audio-visual cues (Swartzentruber, 1995) lends credence to this view, to the extent that stimulus generalisation is likely to be greater within a stimulus modality than between modalities. Such effects could also explain why transfer occurs more readily to occasion-set CSs. Generalisation may occur between occasion setters:  $OS_1$  acting on  $cs_1 \rightarrow US_1$  may transfer more effectively to a  $cs_2$  that has been occasion-set by  $OS_2$  because of generalisation between  $OS_1$  and  $OS_2$ ; if  $cs_2$  has not been occasion-set this source of generalisation is not available. Also generalisation between CSs may not be based solely on physical similarity, but also via their common training history - *acquired equivalence* (Honey & Hall, 1989). For example, Bonardi and Hall (1994a) examined generalisation from occasion-set  $cs_1$  to two further cues  $cs_2$  and  $cs_3$ , where  $cs_2$  had also been the target of occasion setting, but  $cs_3$  had not (Table 2)<sup>5</sup>. After pairing  $cs_1$

---

<sup>5</sup>  $cs_3$  was reinforced and extinguished (Table 2), to give it a similar training history to  $cs_2$  without endowing it with occasion-setting properties.

with food they found greater generalisation of conditioned responding to  $CS_2$  than to  $CS_3$  (Figure 3). They argued that the common training history of the occasion-set cues increased their similarity, which fostered selective transfer between them.

*Table 2 about here*

*Figure 3 about here*

**Summary** Transfer of occasion setting across CSs occurs, but it is typically incomplete. Transfer can be more substantial when the transfer CS had been occasion-set, but also occurs when it has not. These findings support the hierarchical account, which predicts occasion setters should be CS-specific, but that some degree of transfer can occur via stimulus generalisation, either between training and transfer CSs or between occasion setters, and based on stimulus similarity or stimulus training history. These findings are, however, also consistent with configural theory to which we turn next.

### **Configural Theory**

Configural theories were developed to explain performance on nonlinear discriminations such as negative patterning,  $A \rightarrow US / B \rightarrow US / AB \rightarrow \text{nothing}$ . According to the standard associative assumption about summation, a compound stimulus is equivalent to the sum of its parts, so presenting  $A$  and  $B$  together sums their associative strengths. If CR is monotonically related to associative strength, accurate performance is impossible because responding to  $AB$  must be higher than to  $A$  or  $B$  alone. Configural theories abandon this summation assumption, but differ in how they conceptualise the stimulus compound. For example, Rescorla (1972) proposed that  $AB$  comprises the



elements of  $A$  and  $B$  plus a third, configural cue,  $x$  that is only present when  $A$  and  $B$  co-occur; thus negative patterning becomes  $A \rightarrow US / B \rightarrow US / ABx \rightarrow \text{nothing}$ , meaning  $x$  can acquire inhibitory strength and allow solution of the task. In contrast Pearce (1987; 1994) proposed that although  $AB$  is distinct from  $A$  and  $B$ , generalisation can occur between them based on the proportion of elements they share; some have argued this effectively conceptualises the  $AB$  compound as a *subset* of the elements of  $A$  and  $B$  (Brandon, Vogel & Wagner, 2000). Brandon et al. (2000) combined these ideas in the replaced elements model, according to which some elements of  $A$  and  $B$  are replaced by elements unique to  $AB$ . All three can explain negative patterning, because all predict that associative strength accrued to  $A$  and  $B$  is not only source of responding on  $AB$  trials.

Configural theory can thus explain many of the facts about occasion setting by recasting an  $OS_1: cs_1 \rightarrow US_1 / cs_1 \rightarrow \text{nothing}$  discrimination as conditioning of a configural cue  $P$  created by co-occurrence of  $OS_1$  and  $cs_1$ . If what has been conditioned is not  $OS_1$  but  $P$ , manipulations of  $OS_1$ 's associative strength will leave the associative strength of  $P$  relatively intact. Configural accounts can also explain CS specificity effects: in the  $OS_1: cs_1 \rightarrow US_1 / cs_1 \rightarrow \text{nothing}$  discrimination, for example, the configural cue  $P$  is reinforced, and transfer of  $OS_1$  to  $cs_2$ , in an  $OS_1: cs_2$  compound occurs to the extent that there is generalisation between  $P$  and a second configural cue  $Q$  produced by co-occurrence of  $OS_1$  and  $cs_2$ . Many configural theories predict this would depend in part on the similarity of  $cs_1$  and  $cs_2$  — which, as we saw above, is what seems to be the case.

**US-specificity** Both hierarchical and configural theories can thus explain occasion setting and its CS-specificity. But they make different predictions about *US-specificity* - whether  $OS_1$  trained as a signal that  $cs_1$  predicts  $US_1$  ( $OS_1: cs_1 \rightarrow US_1$ ) will act more effectively on other CSs that *also* predict  $US_1$ . The hierarchical account predicts that occasion setters should be US-specific - yet Holland reported no sign of US specificity (perfect transfer) with a CS that predicted  $US_2$ , provided  $US_2$  had also been involved in

an occasion-setting task (Holland, 1989c). However, US specificity has been reported even when  $US_2$  has been occasion-set: Bonardi, Bartle & Jennings (2012) trained rats on two positive-patterning tasks in which  $OS_1$  signalled reinforcement of  $cs_1$  with  $US_1$ , and  $OS_2$  of  $cs_2$  with  $US_2$  (with  $cs_1$ ,  $cs_2$ ,  $OS_1$  and  $OS_2$  also presented alone - Table 3) and then examined the ability of  $OS_1$  and  $OS_2$  to transfer to  $cs_3$  and  $cs_4$ , which predicted  $US_1$  and  $US_2$  respectively. Transfer was greater when the outcomes of the occasion setter and transfer CSs were the *same*, despite the fact that both  $US_1$  and  $US_2$  had been involved in occasion-setting tasks (Figure 4; see also Morell & Davidson, 2002).

*Table 3 about here*

*Figure 4 about here*

These findings are not consistent with the predictions of standard configural theory which assumes that configural cues do not encode information about the USs that they signal. This means they cannot explain how  $OS_1$ , signalling  $cs_1 \rightarrow US_1$ , can transfer more effectively to a stimulus that also signals  $US_1$  than to a stimulus that signals  $US_2$  (Morell & Davidson, 2002; Bonardi et al., 2012). In both cases responding should be determined by the similarity of the trained configural cue  $OS_1:cs_1$  to the  $OS_1:cs_2$  cue present at test, neither of which is affected by the nature of the US. Given the theoretical importance of this issue, we conducted a further study in which 16 rats were trained on two feature-negative discriminations with different USs (Table 4).  $OS_1$  and  $OS_2$  were visual (illumination of two 2.8-W jewel lights pulsed at 1 Hz, or of a 2.8-W bulb mounted inside the food magazine),  $cs_1$  and  $cs_2$  were auditory (white noise or 10-Hz clicker both at 73 dB) and  $US_1$  and  $US_2$  were either 2 sucrose pellets or .3 ml of groundnut oil (for complete description of apparatus see Bonardi et al., 2012). All stimulus presentations were 10-s in duration, and there was a 5-s trace interval between

OS offset and CS onset on compound trials; the intertrial interval (ITI) comprised a fixed 60s plus a variable interval with a mean of 30s, and on reinforced trials US delivery was delivered at CS offset. Each of the first 15 training sessions comprised, in a semi-random order, 8  $OS_1 \rightarrow \text{nothing}$ , 8  $OS_2 \rightarrow \text{nothing}$ , 8  $cs_1 \rightarrow US_1$ , 8  $cs_2 \rightarrow US_1$ , 16  $OS_1: cs_1 \rightarrow \text{nothing}$  and 16  $OS_2: cs_2 \rightarrow \text{nothing}$  trials<sup>6</sup>. In the next 18 sessions the OSs were also reinforced; each of these sessions comprised 4  $OS_1 \rightarrow US_1$ , 4  $OS_2 \rightarrow US_1$ , 4  $cs_1 \rightarrow US_1$ , 4  $cs_2 \rightarrow US_2$  trials, 24  $OS_1: cs_1 \rightarrow \text{nothing}$  and 24  $OS_2: cs_2 \rightarrow \text{nothing}$  trials. The final 6 of these sessions also included three trials with each of two test excitors  $cs_3$  and  $cs_4$ , a 300-Hz buzzer and a 2-kHz tone, both at 75dB, paired with oil and sucrose respectively. These sessions were otherwise identical to those of the preceding stage. The results from training are shown in Figure 5. The rats learned the task, responding more to  $cs_1$  and  $cs_2$  when they were presented alone on reinforced trials than on nonreinforced trials when they were preceded by the OSs; ANOVA with trial type (CS reinforced or not) and session block as factors revealed main effects of trial type,  $F(1, 15) = 6.67$ ,  $MSe = 16.38$ ,  $p = .021$ , block,  $F(10, 150) = 16.36$ ,  $MSe = 8.86$ ,  $p < .001$ , and a significant interaction,  $F(10, 150) = 3.78$ ,  $MSe = 1.82$ ,  $p < .001$ . Simple main effects analysis (using the pooled error term) revealed an effect of trial type on blocks 3, 7, 10 and 11, smallest  $F(1, 15) = 4.62$ ,  $MSe = 16.38$ ,  $p = .048$ . Finally, the mean rate of responding to  $cs_3$  and  $cs_4$  was 11.25 and 12.90 rpm for blocks 10 and 11 respectively.

*Table 4 about here*

*Figure 5 about here*

---

<sup>6</sup> For half the animals the noise was reinforced with sucrose and the click with oil and for the remainder the reverse; for half of each of these subgroups the noise was preceded by the jewel light, and the click by the tray light on compound trials, and for the remainder the reverse.

The ability of  $OS_1$  and  $OS_2$  to suppress responding to  $cs_3$  and  $cs_4$  was then evaluated: in four test sessions rats continued to receive three trials with each test excitator, plus 24 *same* trials in which either  $OS_1$  preceded  $cs_3$ , or  $OS_2$  preceded  $cs_4$ , and 24 *different* trials on which these combinations were reversed (12 of each type). If  $OS_1$ 's effects were specific to  $US_1$  then  $OS_1$  should be more effective on *same* trials with  $cs_3$  that also predicted  $US_1$ , than on *different* trials with  $cs_4$  which predicted  $US_2$ . The test session data are shown in Figure 6 in two, 2-session blocks; ANOVA performed on data from the OS trials with trial type (same or different) and block as factors revealed a main effect of trial type,  $F(1, 15) = 5.77$ ,  $MSE = 2.29$ ,  $p = .03$ , no effect of block,  $F < 1$ , and a significant interaction,  $F(2, 30) = 4.70$ ,  $MSE = .459$ ,  $p = .047$ . Simple main effects analysis revealed an effect of trial type on block 2,  $F(1, 15) = 5.72$ ,  $MSE = 2.24$ ,  $p = .03$ , although not on block 1,  $F(1, 15) = 1.00$ ,  $MSE = 2.24$ ,  $p = .33$ . Thus, just as in the positive patterning task, transfer was greater on *same* trials when the outcomes of the occasion setter and transfer CSs matched, despite the fact that both  $US_1$  and  $US_2$  had been involved in occasion-setting discriminations. Such US specificity cannot be predicted by the configural account, but is perfectly consistent with hierarchical theory.

*Figure 6 about here*

**Specificity to CS/US combination** Another key discriminator between the hierarchical and configural accounts is whether an  $OS_1$  which has signalled  $cs_1 \rightarrow US_1$  and  $cs_2 \rightarrow US_2$  will be equally effective on  $cs_1$  if it is subsequently paired with  $US_2$ , and on  $cs_2$  if it is paired with  $US_1$ . If the occasion setter acts on the actual association between two specific events, as hierarchical theory predicts, then its potency should be substantially reduced if those associations are replaced by different ones: thus an occasion setter should be *specific to a CS/US combination*. In fact, evidence on this is mixed: Holland (1989c) trained rats on two feature-negative tasks with two USs ( $cs_1 \rightarrow US_1 / OS_1$ :

$cs_1 \rightarrow \text{nothing}$  /  $cs_2 \rightarrow US_2$  /  $OS_2: cs_2 \rightarrow \text{nothing}$ ), and then the rats received pairings of  $cs_1$  with  $US_2$ , and  $cs_2$  with  $US_1$ . Transfer of occasion-setting to these new  $cs_1 \rightarrow US_2$ , and  $cs_2 \rightarrow US_1$  associations was complete — response rates to  $cs_1$  signalled by  $OS_1$ , and to  $cs_2$  signalled by  $OS_2$ , were as high after  $cs_1$  and  $cs_2$  retraining as before — despite these event combinations never having been subject to occasion setting. This is inconsistent with the hierarchical account. However, other studies using different techniques have provided support for specificity to particular CS/US combinations. Bonardi & Ward-Robinson (2001) trained pigeons on a *switching* task (Asratyan, 1961) with two USs:  $OS_A$  signalled that  $cs_1$  was followed by  $US_1$ , and  $cs_2$  by  $US_2$ , while  $OS_B$  signalled the opposite (see Table 5).  $OS_A$  and  $OS_B$  were diffuse auditory and visual cues,  $cs_1$  and  $cs_2$  keylights, and  $US_1$  and  $US_2$  were red and white lentils. Then birds received *Same* trials, on which  $OS_A$  and  $OS_B$  signalled the same CS/US relations and a further keylight,  $S$ , was interposed between  $cs_1$  and  $cs_2$  and the USs, and also *Different* trials, which were identical except that  $S$  was replaced by keylight  $D$ , and the original CS/US relations were reversed. At test there was less responding to  $S$  than to  $D$ , a result which can be understood as a type of blocking. If  $OS_A$  and  $OS_B$  modulated specific CS/US pairings, on *Same* trials this would ensure that US delivery was fully predicted by  $cs_1$  and  $cs_2$ , and thus that acquisition of associative strength by  $S$  would be blocked. But on *Different* trials  $A$  and  $B$  would not allow  $cs_1$  and  $cs_2$  to predict the outcomes that were delivered, making them surprising and thus able to support learning about  $D$ .

*Table 5 about here*

Bonardi (2007) reported analogous results in a feature-negative task: rats were trained that  $cs_1$  and  $cs_2$  were reinforced with  $US_1$  and  $US_2$  respectively, except when they were signalled by  $OS$  (see Table 6).  $cs_1$  and  $cs_2$  were auditory,  $US_1$  and  $US_2$  oil and sucrose, and  $OS$  visual. Then the rats were trained on two feature-positive tasks with  $OS$

as the feature. In Group Same *OS* signalled that  $cs_1$  was followed by  $US_1$ , and  $cs_2$  by  $US_2$ , but in Group Different these pairings were reversed. If in stage 1 animals learned that *OS* signalled specific  $cs_1 \rightarrow$  no  $US_1$  and  $cs_2 \rightarrow$  no  $US_2$  associations, then learning should be more difficult in Group Same, who unlike Group Different had to learn exactly the *opposite* of what they had learned in stage 2. That is what was observed (Figure 7).

*Table 6 about here*

*Figure 7 about here*

**Perceptual Interactions** Another factor that discriminates between hierarchical and configural theory relates to the conditions fostering occasion-setting. Configural accounts assume that configural cues form most readily when there is a possibility for perceptual interaction between the to-be-configured cues - yet the conditions promoting formation of occasion setters are often not those that would be likely to facilitate such interaction (Holland, 1992). For example, configuring in a feature-positive task seems more likely when feature and target are presented *simultaneously* rather than serially. Yet Holland (1989b) demonstrated that in a feature-positive task with simultaneous OS and CS presentation ( $OS \ \& \ cs \rightarrow US / cs \rightarrow nothing$ ), the OS behaved more like a simple Pavlovian CS; its ability to promote responding to the CS transferred well to other CSs and was attenuated by counterconditioning. But if the OS *preceded* the CS during training ( $OS \rightarrow cs \rightarrow US / cs \rightarrow nothing$ ), it did not transfer to other cues and was not affected by counterconditioning - it was more like an occasion setter. Occasion setting was thus fostered by serial training procedures that were less likely to support configural learning. Configuring also seems more likely when OS and CS are of the same modality as this would allow perception of the two cues to interact more effectively. Yet Holland (1989b) found that simultaneous positive-patterning tasks ( $OS \ \& \ cs \rightarrow US / cs \rightarrow nothing /$

$OS \rightarrow nothing$ ) were easier to learn when OS and CS were of the *same* modality, whereas serial positive-patterning tasks ( $OS \rightarrow cs \rightarrow US / cs \rightarrow nothing / OS \rightarrow nothing$ ) were easier if OS and CS were of *different* modalities. Thus the serial tasks that foster occasion setting were easier to learn if configuring was *less*, rather than more, likely. Conversely, the simultaneous tasks that were *less* likely to result in occasion setting were easier to learn when configuring was more likely.

Other evidence from our laboratory, which addressed the ability of occasion setters to signal *when* a US is delivered (Bonardi & Jennings, 2007), casts further doubt on this perceptual interaction view. Rats were trained on a switching task with two occasion setters,  $OS_1$  and  $OS_2$ , and two target stimuli  $cs_1$  and  $cs_2$ .  $OS_1$  and  $OS_2$  were 10s in duration, and followed by  $cs_1$  or  $cs_2$ , OS and CS being separated by a 5-s trace interval;  $cs_1$  and  $cs_2$  were both presented in *long* and *short* trials, 30-s and 6-s in duration respectively, giving four trial types,  $cs_1$ -*short*,  $cs_1$ -*long*,  $cs_2$ -*short*,  $cs_2$ -*long*. All CS presentations were immediately followed by food, but the OSs gave information about the delay to food delivery:  $OS_1$  signalled  $cs_1$ -*short* and  $cs_2$ -*long* trials, while  $OS_2$  signalled  $cs_1$ -*long* and  $cs_2$ -*short* trials ( $OS_1: cs_1$ -short  $\rightarrow US / OS_1: cs_2$ -long  $\rightarrow US / OS_2: cs_1$ -long  $\rightarrow US / OS_2: cs_2$ -short  $\rightarrow US$ ). Responding on probe trials (identical to training trials except that both  $cs_1$  and  $cs_2$  were presented for 90s and no US was delivered) indicated the rats had learned this task: they showed a peak of responding at around 6s after CS onset on short trials, and 30s after CS onset on long trials. These results could be compatible with a configural account if the trace of each OS and CS decayed in some time-dependent manner after its onset. Thus the rats might learn that on long trials on which  $OS_1: cs_2$  and  $OS_2: cs_1$  were reinforced, a configural cue of the trace of the OS 35s after its offset *plus* the trace of the CS 30s after its offset, was paired with food, and could control timed responding. We reasoned that these temporally sensitive configural cues would be disrupted if we altered the trace interval between the feature and the

target. Thus in a subsequent test we compared responding on training trials with that on trials in which the interval between OS offset and CS onset was increased from 5 to 29s. The rats continued to respond at the 'correct' time after CS onset - despite the fact that the OS trace at this point would have been completely different to that present during training. This is more consistent with a hierarchical-type account which allows each OS to signal a specific temporal relationship with a specific CS and outcome delivery.

**Summary** Occasion setters are not only specific to the CS, but also to the US, with which they were trained. As to whether the occasion setter acts on a specific combination of a CS and US, the evidence is mixed; while transfer studies do not support this prediction, other types of task do. In combination with findings from operant tasks examining the specificity of discriminative stimuli and inhibitors to specific response-reinforcer associations (Colwill & Rescorla, 1990; Colwill, 1991; Bonardi & Hall, 1994b), these data confirm the view that occasion setters are *association-specific* in their action. This pattern of findings is not anticipated by configural theory, but is consistent with the hierarchical account. A further problem for configural theory stems from its implicit assumption that configuring is more likely when the to-be-configured cues may interact perceptually. In fact it seems that the more likely such interaction is, the less likely occasion setting is to occur. Also occasion-setting-like behaviour can be maintained even when the potential configural cues are severely degraded. Thus, although the parsimony of a configural account is appealing, standard versions of such accounts have difficulty dealing with the empirical findings.

### **Mechanisms of Occasion-Setter Formation**

In comparison with the body of work on transfer, relatively few studies have explored the learning process by which occasion setting is established. Configural



accounts predict that formation of occasion setters should obey normal associative rules, provided the to-be-associated stimulus is the configural cue of OS and CS. But because the hierarchical account assumes that occasion setters act non-associatively, there is no reason why their formation should be governed by associative rules - and it is not clear what the alternative should be.<sup>7</sup> One starting point was suggested by Mackintosh, who interpreted Skinner's original formulation by suggesting that the OS (or an operant  $S^d$ ) controls operation of  $CS \rightarrow US$  or  $(R \rightarrow US)$  through an associative-type process: "*If the  $S^d$  provides subjects with information about the relationship between their actions and consequences, this is because it is associated not with those actions, nor with their consequences, but with a representation of the relationship between them.*" (Mackintosh, 1983; pp.110-111). Thus although the OS's effect on behaviour is, by definition, not mediated by an association with either CS or US, its ability to exert this control is nonetheless the product of an association with the 'relationship' between them. Although this notion has persisted (e.g., Bonardi, 1996; Bonardi et al., 2012), it is not well specified; nonetheless, the assumptions on which it relies can be tested. For example, it predicts that the process by which an occasion setter forms should obey standard associative rules. Thus, assuming that the 'event' to be associated with the OS is the  $CS \rightarrow US$  pairing, acquisition of occasion setting should show blocking. A corollary of this is that the  $CS \rightarrow US$  [association](#) can be regarded as a unitary 'event' that can enter into associations. We will consider evidence relating to both of these proposals.

**Learning rules:** Bonardi (1991) trained rats that during 3-min presentations of  $OS_A$ , 5-sec presentations of  $cs_I$  were followed by a food US; in  $OS_A$ 's absence  $cs_I$  was nonreinforced ( $OS_A: cs_I \rightarrow US / cs_I \rightarrow nothing$ ). In stage 2 animals were still trained with  $OS_A$  and  $cs_I$ , but  $OS_A$  was presented in compound with  $OS_B$  ( $OS_A OS_B: cs_I \rightarrow US /$

---

<sup>7</sup> the exception is Rescorla's US modulation account; in a parallel with conditioned inhibition, he proposed that occasion setters were established in the presence of the reinforcement of a CS that possessed inhibitory strength.

$cs_1 \rightarrow \text{nothing}$ );  $OS_B$  in compound with  $OS_C$  also signalled reinforcement of  $cs_2$  ( $OS_B OS_C$ :  $cs_2 \rightarrow US / cs_2 \rightarrow \text{nothing}$ ). Finally the ability of  $OS_B$  to elevate responding to  $cs_1$  and  $cs_2$  was evaluated. If in stage 1  $OS_A$  became associated with  $cs_1 \rightarrow US$ , this would block formation of an association between  $OS_B$  and  $cs_1 \rightarrow US$  in stage 2; but as  $OS_C$  was novel,  $cs_2 \rightarrow US$  pairings would not be predicted by  $OS_B OS_C$ , and  $OS_B$ 's association with  $cs_2 \rightarrow US$  would be unimpaired. If  $OS_B$ 's ability to act as an occasion setter, promoting responding to  $cs_1$  and  $cs_2$ , was based on its association with  $cs_1 \rightarrow US$  and  $cs_2 \rightarrow US$ , then  $OS_B$  should elevate responding to  $cs_2$  more than to  $cs_1$  - which is what was observed. Critically this cannot be due to blocking of the Pavlovian association between  $OS_B$  and food, as this would affect responding to  $cs_1$  and  $cs_2$  equally. Configural theory could also explain these results via blocking. Generalisation between  $OS_A$ :  $cs_1$  and  $OS_A OS_B$ :  $cs_1$ , could curtail acquisition of associative strength by  $OS_A OS_B$ :  $cs_1$ , reducing generalisation to  $OS_B$ :  $cs_1$ , which would thus elicit less responding at test than  $OS_B$ :  $cs_2$

Related evidence has been generated in operant tasks. Colwill & Rescorla (1990) trained rats on two operant contingencies signalled by  $OS_x$  ( $OS_x$ :  $R_1 \rightarrow US_1 / R_1 \rightarrow \text{nothing} / OS_x : R_2 \rightarrow US_2 / R_2 \rightarrow \text{nothing}$ ). In further training  $OS_x$  was accompanied by either  $S$  or  $D$ . During  $OS_x$  &  $S$  trials the same operant contingencies were in operation as before ( $OS_x$  &  $S$ :  $R_1 \rightarrow US_1 / R_1 \rightarrow \text{nothing} / OS_x$  &  $S$ :  $R_2 \rightarrow US_2 / R_2 \rightarrow \text{nothing}$ ) whereas during  $OS_x$  &  $D$  trials each response was paired with the alternative outcome ( $OS_x$  &  $D$ :  $R_1 \rightarrow US_2 / R_1 \rightarrow \text{nothing} / OS_x$  &  $D$ :  $R_2 \rightarrow US_1 / R_2 \rightarrow \text{nothing}$ ). At test  $D$  was better able than  $S$  to elicit  $R_1$  and  $R_2$ . This is interpretable as blocking by the *occasion setter*. Initially  $OS_x$  becomes associated with  $R_1 \rightarrow US_1$  and  $R_2 \rightarrow US_2$ . In stage 2, during  $OS_x$  &  $S$ ,  $R_1 \rightarrow US_1$  and  $R_2 \rightarrow US_2$  are fully predicted by  $OS_x$ , blocking the ability of  $S$  to become associated with them. But when  $R_1$  and  $R_2$  are paired with the alternative outcome during  $OS_x$  &  $D$ , the resultant contingencies are surprising, so  $D$  may become associated with them. Again configural theory would treat this as a case of unblocking - in stage 2 the

outcomes following  $OS_x$  &  $S$  and  $R_I$ , for example, are better predicted than those following  $OS_x$  &  $D$  and  $R_I$ .

**Summary** Occasion-setter formation appears to conform to associative rules, as it shows blocking: the ability of  $OS_I$  to acquire control over  $CS_I \rightarrow US_I$  is blocked if a pretrained occasion setter for  $CS_I \rightarrow US_I$  is present. For the reasons we have outlined, this is not predicted by the hierarchical account unless it assumes that (i) each CS/US pairing may act as a unitary outcome independent of its constituents, and (ii)  $OS$ s acquire their occasion-setting properties through becoming associated with the CS/US pairing.

### **Elaborations of Hierarchical and Configural Theory**

**Elaborated Hierarchical Account** Incorporation of hidden units into the hierarchical structure can help accommodate these assumptions. Bonardi et al. (2012) suggested an elaboration of the theory which assumes that when a  $CS \rightarrow US$  association is established, a hidden unit is recruited that is specific to that association and linked to it via an asymmetric link. When the association is active, the hidden unit is activated in a normal associative manner (Figure 8 top panel; upward arrow), and any other stimulus that is present becomes associated with the hidden unit via standard associative learning; conversely activation of the hidden unit gates the ability of the CS to predict the US (Figure 8 top panel; round-headed arrow). Thus a cue that is present when the CS-US association is established becomes associated with its hidden unit, allowing presentation of that cue to facilitate flow of activation from CS to US, and act as an occasion setter<sup>8</sup>.

---

<sup>8</sup> One possibility is that during  $CS \rightarrow US$  formation, when the US is surprising, the hidden unit is recruited into  $A1$  (cf. Wagner, 1980), and can support learning both as a 'CS' and as a 'US'; once the  $CS \rightarrow US$  association is formed, and the CS successfully predicts the US, the hidden unit enters  $A2$ , and can support learning as a CS, not a US.

*Figure 8 about here*

This modified theory can explain blocking of occasion setting (Bonardi, 1991): in stage 1 an association linking  $OS_A$  to the hidden unit modulating  $cs_1 \rightarrow US$  forms and reaches asymptote. In stage 2 this hidden unit is now fully predicted by  $OS_A$ , so formation of an association between the added  $OS_B$  and the hidden unit is blocked, and  $OS_B$  cannot acquire any occasion-setting properties for  $cs_1 \rightarrow US$ . In contrast, when  $OS_B$  and  $OS_C$  signal the new link  $cs_2 \rightarrow US$ , a new hidden unit is recruited, with which both  $OS_B$  and  $OS_C$  can become associated. Thus  $OS_B$  will control responding to  $cs_2$  but not  $cs_1$ .

**Nature of the association representation** A second corollary of this modified hierarchical account relates to whether a particular *combination* of events can enter into further associations. Evidence on this point has emerged from studies of acquired equivalence - the observation that the functional similarity between two cues can be enhanced if they share a common training history. Honey and Hall (1989) paired stimulus *A* with food, and then subsequently with shock. They observed that the fear CR elicited by *A* generalised more readily to a second stimulus *B* if it had also been paired with food. They attributed this to mediated conditioning (Holland, 1981; Ward-Robinson & Hall, 1999): when *A* was paired with shock the representation of the food, with which it had previously been associated, was activated and became associated with the shock via a mediated conditioning process. Thus *B*, also able to activate the food representation, could also elicit fear at test via this food  $\rightarrow$  shock association. This logic has been applied to occasion-setting. Honey and Watt (e.g. 1998; cf. Bonardi & Jennings, 2009) trained rats on a switching task with four, 10-s auditory occasion setters,  $OS_1$ ,  $OS_2$ ,  $OS_3$  and  $OS_4$ , and two 10-s visual CSs,  $cs_1$  and  $cs_2$ . When preceded by either  $OS_1$  or  $OS_3$ ,  $cs_1$  was reinforced with food and  $cs_2$  was not, but when preceded by  $OS_2$  or  $OS_4$  these contingencies were reversed (see Table 7). Thus  $OS_1$  and  $OS_3$  signalled the same

contingencies between  $CS_1$ ,  $CS_2$  and food as did  $OS_2$  and  $OS_4$ . Then  $OS_1$  was paired with shock while  $OS_2$  was nonreinforced, and finally fear of  $OS_3$  and  $OS_4$  was evaluated. In a parallel of the acquired equivalence result described above, they found greater fear of  $OS_3$ , trained in the same manner as the shock-reinforced  $OS_1$ , than  $OS_4$ .

*Figure 9 about here*

These results can also be explained by the elaborated hierarchical account, as it predicts that the hidden units relating to the  $CS_1 \rightarrow US_{food}$  and  $CS_2 \rightarrow nothing$  pairings will become activated during the shock conditioning stage. During initial training both  $OS_1$  and  $OS_3$  become associated with the hidden units linked to the  $CS_1 \rightarrow US_{food}$  (Figure 9) and  $CS_2 \rightarrow nothing$  links. When  $OS_1$  is then paired with shock, the hidden units linked to both  $CS_1 \rightarrow US_{food}$  and  $CS_2 \rightarrow nothing$  are activated and become associated with shock<sup>9</sup>. When  $OS_3$  is presented at test, as it can also activate these  $CS_1 \rightarrow US_{food}$  and  $CS_2 \rightarrow nothing$  hidden units, it also indirectly activates the representation of shock and elicits fear.

This implies that the  $CS_1 \rightarrow US_{food}$  and  $CS_2 \rightarrow nothing$  pairings should themselves become aversive as a result of this training, as use of these associations will activate their respective hidden units, and hence the shock representation. This prediction was tested in two further studies (Bonardi & Jennings, 2009). Rats were given the same training as in Honey & Watt's (1998) study, but in a final test fear of the  $CS_1 \rightarrow US_{food}$  and  $CS_2 \rightarrow nothing$  pairings was evaluated: rats were trained to respond for food on an operant baseline, and while they were responding they experienced presentations of  $CS_1$  and  $CS_2$ ; in one study both were paired with food, and in the other both were nonreinforced. The rate of operant responding *after* each of these pairings was compared to baseline rates of responding before CS onset. If the above analysis is correct, animals should show more

---

<sup>9</sup> In terms of the suggestion made above, if use of the association recruits the hidden unit into A2, it can enter further learning as a CS, and thus become associated with shock.

fear after  $cs_1 \rightarrow US_{food}$  than after  $cs_2 \rightarrow US_{food}$  trials, but conversely more fear after  $cs_2 \rightarrow nothing$  than  $cs_1 \rightarrow nothing$  trials. This was what was found (Figure 10).

*Table 7 about here*

*Figure 10 about here*

**Extended Configural Theory** This elaboration of the hierarchical account has allowed it to expand its range of predictions, and a similar strategy was [applied to](#) configural theory. Honey and Watt proposed an 'extended' configural theory (e.g. Honey & Watt, 1999; Honey & Ward-Robinson, 2002) to allow encoding of the US representation, and hence account for US specificity effects. It assumes that configuring occurs not through perceptual interaction, but via hidden units (e.g. Pearce, 1994). An associative structure is supposed comprising input and output units with an intervening set of hidden units. A given CS or US will uniquely correspond to a specific input or output unit respectively, but CS/US pairings can also recruit a hidden unit that links their input and output units (Figure 8 bottom panel). This effectively allows encoding of the US in the configural cue: for example, if  $OS_1$  signals that  $cs_1$  predicts  $US_1$ ,  $OS_1$  and  $cs_1$  will become associated with the same hidden unit  $p$  that is also linked to the  $US_1$  outcome (Figure 11a). This explains why  $OS_1$  might show US-specificity, and transfer better to  $cs_2$  if it is associated with  $US_1$  than with  $US_2$ . Responding will occur to the extent that the  $OS_1: cs_2$  compound can activate a hidden unit; thus if  $cs_2$  is paired with  $US_1$  it will also be linked to the critical hidden unit  $p$ , and elicit more responding than if it were paired with  $US_2$  and thus linked to a second hidden unit  $q$ . Note that this prediction assumes that two sources of activation to  $p$  have a greater effect than one source of activation to each of  $p$  and  $q$  (Figure 11a (i) and (ii)). If such hidden units are recruited during normal conditioning with a single CS, then this leads to the prediction that responding should

*always* be greater to a compound of two CSs that predict the same US than when they predict different USs. Evidence on this is mixed, however: some studies find more Pavlovian summation when the predicted USs match (Rescorla, 1999), others more responding when they differ (Watt & Honey, 1997), and others no difference (Ganesan & Pearce, 1988). Nor can this account clearly predict CS/US specificity: if  $OS_1$  signals  $cs_1 \rightarrow US_1$ , and  $OS_2$  signals  $cs_2 \rightarrow US_2$ , hidden units  $p$  and  $q$  form (Figure 11b). If  $cs_1$  were then paired with  $US_2$ , and  $cs_2$  with  $US_1$  (dotted arrows) the model would presumably predict that *both* CSs will become associated with *both* hidden units; thus without further assumptions there is no reason to predict that  $OS_1$ :  $cs_1$  would be any different in its effect on behaviour than  $OS_1$ :  $cs_2$ .

*Figures 11a and 11b about here*

Because it [is an](#) associative model, the extended configural account can explain why occasion-setter formation obeys associative-type principles, provided the rules governing association of input and hidden units are subject to standard cue competition. For example, blocking of occasion setting could be explained (Bonardi, 1991): if  $OS_A$  and  $cs_1$  are linked to the same hidden unit as the  $US$ , then as long as the same hidden unit is recruited in both stages of training,  $OS_A$  will be able to block acquisition of associative strength by  $OS_B$  for this hidden unit, impairing  $OS_B$ 's ability to become an occasion setter. The model could also explain the results reported by Honey and Watt (1998) and Bonardi & Jennings (2009; Figure 12). Because both  $OS_1$  and  $OS_3$  signal that  $cs_1$  predicts  $US_{food}$ , all these events will become associated with  $p$ , and so on. When  $OS_1$  is subsequently paired with shock,  $p$  is activated and also becomes associated with shock; thus anything that can also activate  $p$  - such as  $OS_3$  - will also access the shock representation and elicit fear. However, explaining Bonardi & Jennings' (2009) results requires the additional assumption that presentation of  $US_{food}$  can also activate the

hidden unit via a feedback mechanism (Figure 12, dotted arrows). Thus presentation of  $cs_1$  and  $US_{food}$  provides two sources of activation to the  $p$  hidden unit associated with  $US_{shock}$ , while presentations of  $cs_2$  and  $US_{food}$  each activate only one hidden unit that is linked to  $OS_1$  ( $p$  and  $q$ ); thus as two sources of activation to one hidden unit have a greater effect than one source of activation to two different hidden units, presentations of  $cs_1$  and  $US_{food}$  will produce more fear than presentations of  $cs_2$  and  $US_{food}$ .

*Figure 12 about here*

**Summary** Extended configural theory overcomes the inability of more traditional configural theories to explain US specificity, but at some expense. Current formulations do not specify the conditions required for recruitment of hidden units, and the account must also make additional assumptions - for example, about the effect of summing two sources of activation to one hidden unit, or how US presentation can also activate its associated hidden unit. Elaborated hierarchical theory must also make assumptions, about the properties of the associations linking the association to its hidden unit, the rules governing how the hidden unit is activated and when it can undergo learning. Given this potential impasse, a different means is required to discriminate between these approaches. We now turn to some different classes of evidence that might achieve this.

### **Further Discriminating Evidence**

**OS/CS symmetry** One fundamental distinction between the elaborated hierarchical and extended configural theories lies in the role of the occasion setter. Configural theories assume only the processes of associative learning: activity in a stimulus representation can only influence activation of a second stimulus via an association between them. Thus there is no qualitative distinction between the action of the OS and CS, because



there is only one way that one stimulus can interact with another<sup>10</sup>. In contrast hierarchical theory proposes that CS representations can also interact in a *nonassociative* manner, that allows an occasion setter to facilitate the flow of activation in an associative link. Thus the hierarchical account assumes that the OS plays a *qualitatively distinct* role from that of a CS in controlling behaviour.

We examined this distinction in two studies (Bonardi et al., 2012). Rats were trained on two positive-patterning discriminations with two occasion setters  $OS_1$  and  $OS_2$ , two conditioned stimuli  $cs_1$  and  $cs_2$  and two USs  $US_1$  and  $US_2$  (Table 8). According to extended configural theory this should result in formation of two hidden units, one linking  $OS_1$  and  $cs_1$  to  $US_1$ , and another linking  $OS_2$  and  $cs_2$  to  $US_2$  (cf. black arrows Figure 11b). The OSs were visual, the CSs auditory, and the USs sucrose and oil. Stimulus presentations were of 10-s duration, and OS offset was separated from CS onset by a 5-s trace interval. The rats were also trained with two types of transfer stimulus (i) two separately trained *test excitors*  $cs_3$  and  $cs_4$  which, like the CSs, were auditory and were immediately followed by either  $US_1$  or  $US_2$ ; and (ii) two separately trained *pseudo-occasion setters*  $POS_1$  and  $POS_2$  which, like the OSs, were visual and were paired with either  $US_1$  or  $US_2$  after a 15-sec trace interval.

*Table 8 about here*

*Figure 13 about here*

The studies differed in their test procedure. In one we examined performance on trials that were identical to the compound training trials except that  $cs_1$  and  $cs_2$  were replaced by one of the test excitors,  $cs_3$  or  $cs_4$ . The constituents of these test compounds

---

<sup>10</sup> Although earlier theories, as we saw above, allowed for inhibitors to modulate activation of representations rather than activating anything directly, later authors have rejected this possibility (Mackintosh, 1983).

either both signalled the *same* outcome ( $OS_1: cs_3 / OS_2: cs_4$ ) or *different* outcomes ( $OS_1: cs_4 / OS_2: cs_3$ ). Consistent with the US specificity predicted by both elaborated hierarchical and extended configural theory, we found more responding to  $cs_3$  and  $cs_4$  on *same* than on *different* trials. In the second study the test trials were identical to the compound training trials except this time  $OS_1$  and  $OS_2$  were replaced by one of the pseudo-occasion setters  $POS_1$  and  $POS_2$ . These trials were again classified as either *same* ( $POS_1: cs_1 / POS_2: cs_2$ ) or *different*, ( $POS_1: cs_2 / POS_2: cs_1$ ). According to extended configural theory, as there is no qualitative distinction in the ability of OS and CS to activate the hidden unit, replacing either element of the trained compound will produce the same effect, more responding on *same* than *different* trials; but this was not observed. Mean rates of responding on *same* ( $s$ ) and *different* ( $d$ ) trials in each of the 2-trial blocks of each test were converted to a ratio of form  $s/d$  and averaged; ratios greater than 1 indicate more responding on same than different trials. The data, in Figure 13, suggest that the ratios exceeded 1 in the first study ( $p = 0.27$ ), but not in the second; ANOVA revealed a significant effect of experiment,  $F(1,30) = 4.74$ ,  $MSe = .493$ ,  $p = .04$ .

We interpreted these findings as supporting the hierarchical view that OS and CS do not play functionally equivalent roles. Although the CS acts in a standard associative manner, the OS does not. Thus replacing the CS of the training compound with a separately trained stimulus retains the occasion-setting nature of the task in a way that replacing the OS with a separately **trained** pseudo-occasion setter does not, meaning the hierarchical account need not predict the same pattern of responding in both tests. In contrast, the extended configural theory must predict that both tests are functionally the same: both  $OS_1$  and  $CS_1$  are linked to the same hidden unit, and so replacing either of these cues should produce more responding if the replacement cue is associated with the same US as the cue that remains. Moreover, we noted above that this prediction relies on specific assumptions about the activation thresholds for the hidden units, which means it has difficulty explaining the finding that summation is greater when the pair of CSs

predict two different USs than when they predict the same US (Watt & Honey, 1997). It is thus of interest that in the final study of the series we combined simple CSs predicting the same or different outcomes, and found *less* responding when the USs matched<sup>11</sup>. A dissociation in the pattern of responding to CS compounds whose components predict the same, or different, outcomes depending on whether one of those components is an occasion setter does not support the extended configural account.

**Parsing of the Association** Another study attempted to discriminate the two theories in a different way. In a task similar to that described above two groups of 8 rats received training trials in which  $cs_1$  and  $cs_2$ , 10-sec presentations of a 74-dB, 10-Hz clicker, or the illumination of the tray light, were each paired with one of two USs; for half the rats the click was paired with sucrose and the traylight with oil, and for the rest the reverse. They also received nonreinforced compound trials on which a third cue  $OS$ , the 10-s illumination of a 2.8-W jewel light, signalled simultaneous presentation of  $cs_1$  and  $cs_2$  after a 5-s trace interval ( $OS: cs_1 \& cs_2 \rightarrow nothing / cs_1 \rightarrow US_1 / cs_2 \rightarrow US_2$ ).<sup>12</sup> The data from the discrimination training phase are presented in Figure 14; ANOVA showed that both groups learned the discrimination,  $F(1, 14) = 17.82, p = .001$ ; smallest  $p$  involving the group factor:  $F(1, 14) = 1.06, p = .32$ . Then, to equate pairings of  $cs_1$  and  $cs_2$  with the two USs before the test, rats received *reverse* training in which  $cs_1$  was paired with  $US_2$  and  $cs_2$  with  $US_1$ <sup>13</sup>. The groups did not differ in this phase, and responding remained high (20.1 and 24.9 rpm for Groups Same and Different respectively on the final 2-session block; see below); ANOVA with group and 2-session block as factors revealed

---

<sup>11</sup> We tested the effect of signalling  $cs_1$  and  $cs_2$  with  $cs_3$  and  $cs_4$ , again either in *same* ( $cs_3 \rightarrow cs_1$  and  $cs_4 \rightarrow cs_2$ ) or *different* ( $cs_3 \rightarrow cs_2$  and  $cs_4 \rightarrow cs_1$ ) combinations, and found significantly higher responding on different trials in the first test block,  $p = 0.24$ .

<sup>12</sup> There were 56 trials scheduled per session, with a variable inter-trial interval (mean of 75s, range of 60-90s), and 26 training sessions in this stage; each comprised 4  $cs_1 \rightarrow US_1$  and 4  $cs_2 \rightarrow US_2$  trials, and in all but the first 10 sessions also 48  $OS: cs_1 \& cs_2 \rightarrow nothing$  trials.

<sup>13</sup> This training comprised 18 further sessions, each of 8  $cs_1 \rightarrow US_2$  and 8  $cs_2 \rightarrow US_1$  trials.

nothing significant; the smallest  $p$  was associated with the interaction,  $F(8, 112) = 1.07$ ,  $p = .07$ , and the effect of group was not significant,  $F < 1$ .

*Figure 14 about here*

*Table 9 about here*

Finally the rats were divided into two groups, and trained on two feature-positive discriminations. In Group Same the OS signalled reinforcement of  $cs_1$  with  $US_1$ , and  $cs_2$  with  $US_2$ , the opposite of what had been the case during the initial feature-negative training when OS signalled  $cs_1$  would *not* be followed by  $US_1$ , and  $cs_2$  would *not* be followed by  $US_2$ . In contrast, for Group Different the CS→US pairings during feature-positive training were reversed, so test training did not directly contradict the feature-negative discrimination in this way<sup>14</sup>.

The extended hierarchical account assumes that during feature-negative training,  $cs_1 \rightarrow US_1$  and  $cs_2 \rightarrow US_2$  associations form, and that one hidden unit is recruited for each. Although it does not specify how negative occasion setters form, let us assume for simplicity that a stimulus that is present when  $cs_1$  is paired with the unexpected omission of  $US_1$  acquires modulatory control over a  $cs_1 \rightarrow noUS_1$  association, which reduces the extent to which  $cs_1$  can predict  $US_1$ . Thus the OS would acquire inhibitory control over the  $cs_1 \rightarrow US_1$  and  $cs_2 \rightarrow US_2$  associations, and this control would not be compromised when different  $cs_1 \rightarrow US_2$  and  $cs_2 \rightarrow US_1$  associations were subsequently formed. Thus Group Same, who had to learn the opposite of the relations in operation during feature-

---

<sup>14</sup> There were two phases in this stage; in the first, in which the feature-positive discriminations were acquired, comprised 10 sessions, each of 24 OS → nothing, 12  $cs_1 \rightarrow nothing$ , 12  $cs_2 \rightarrow nothing$ , and for Group Same 4 OS:  $cs_1 \rightarrow US_1$  4 OS:  $cs_2 \rightarrow US_2$  trials; the second, in which performance was tested, comprised 12 sessions identical to those of the first phase except that there could be 2, 4 or 8 OS:  $cs_1 \rightarrow US_1$  and OS:  $cs_2 \rightarrow US_2$  trials per session. For Group Different  $US_1$  and  $US_2$  were reversed.

negative training at test, would perform less effectively at test than Group *Different*, as only the former group would have to overcome the inhibitory modulation over these specific associations that was established in the feature-negative training stage.

Extended configural theory instead predicts that initial negative-patterning training would result in hidden unit  $p$  linking  $cs_1$  and  $US_1$ , and  $q$  linking  $cs_2$  and  $US_2$ . The model assumes that learning about non-reinforced trials is governed by an excitatory association from the hidden unit to a no-US representation (e.g., Allman, Ward-Robinson & Honey, 2004; cf. Konorski, 1967). Thus  $OS$ ,  $cs_1$  and  $cs_2$  should be equally associated with a no- $US_1$  representation via hidden unit  $r$ , and a no- $US_2$  representation via the hidden unit  $s$  (Figure 15). When  $cs_1$  and  $cs_2$  are each paired with the alternative USs, they will become linked to  $q$  and  $p$  respectively, resulting in the associative structure shown in Figure 15. It is clear from the symmetry here that there would be no good reason to predict why animals in Group Same should learn their feature-positive test discrimination any less effectively than those in Group Different.

*Figure 15 about here*

The results of this experiment may be seen in Figure 16. Levels of responding declined gently over the course of testing, presumably because the number of reinforced trials per session was on average lower than it had been in the previous phases (see footnote 12). While Group *Different* performed accurately, responding more on reinforced trials on which  $cs_1$  and  $cs_2$  were signalled by  $OS$  than when  $cs_1$  and  $cs_2$  were presented alone, Group *Same* showed the opposite pattern. ANOVA with group, trial type (reinforced and nonreinforced) and blocks as factors revealed a significant interaction between group and trial type,  $F(1, 14) = 7.47, p = .02$ , and simple main effects revealed that the effect of trial type was significant in Group Different,  $F(1, 14) =$

6.37,  $p = .02$ , but not in Group Same,  $F(1, 14) = 1.80$ ,  $p = .20$ . The results were, therefore, in accord with the predictions of the hierarchical account.<sup>15</sup>

*Figure 16 about here*

### **Conclusions: Do We Need a Hierarchical Theory?**

We have reviewed theories of occasion setting which fall into two classes, those that require modification of existing associative learning principles, and those that do not. In the former category the US modulation and memory systems accounts were largely unsupported by evidence on the CS- and US-specificity of occasion setting, leaving the hierarchical account as the main contender. Configural theories fall into the second class, but are unable to explain why occasion setters are US-specific without considerable modification, resulting in extended configural theory. Conversely, hierarchical theory could not explain the apparent constraints on occasion-setter formation without elaboration of its associative structure. Both theories must therefore make a number of assumptions to account for the existing data, and more precise specification of the conditions required for recruitment of hidden units is required, especially for extended configural theory. Future work could usefully compare the degree to which these two relatively informal approaches can be expressed as computational models, and explore any concrete predictions that are generated in this way. Equally, the extent to which the more formal models that have been developed (cf.

---

<sup>15</sup> It should be noted that there is an inevitable confound in this experiment, as the  $cs_1 \rightarrow US_1$  and  $cs_2 \rightarrow US_2$  associations were always learned before the  $cs_1 \rightarrow US_2$  and  $cs_2 \rightarrow US_1$  associations. Thus the occasion-set relations of the test were learned first for Group Same but second for Group Different, and this could also be responsible for the differences at test. However, if the reverse-order training preceded the negative occasion-setting training there would be no guarantee that OS acquired negative occasion-setting properties over  $cs_1 \rightarrow US_1$  and  $cs_2 \rightarrow US_2$  and not to some extent over  $cs_1 \rightarrow US_2$  and  $cs_2 \rightarrow US_1$  - fatally compromising the logic of the experiment.

Kutlu & Schmajuk, 2012; Vogel, Ponce & Wagner, this volume) can accommodate the key classes of evidence described here could be explored. An alternative means of discriminating the two approaches could focus on their underlying assumptions - the theories differ fundamentally in how they conceptualise the roles of OSs and CSs - and this may provide grounds for discriminating between them. We reported two of our own studies that address this distinction, and the results were consistent with hierarchical theory's predictions. It is also possible that neuroscientific approaches may be able to dissociate the two mechanisms.

Of course these two approaches should not be taken as mutually exclusive. There is evidence outside the occasion-setting literature that supports the existence of configural cues - although whether such evidence can be accommodated within the perceptual interaction version of configural theory, or requires the extended configural interpretation, is unclear. The question is thus more one of whether an additional notion of hierarchical control is also required. We have argued in this article that it is, and that any general associative theory that is able to explain occasion-setting effects will need to incorporate such a mechanism. In our view the development of properly formalised mechanisms for hierarchical control, and their incorporation into a more general model, probably represent the path to a truly comprehensive account of associative and occasion-setting phenomena.

#### Acknowledgements

This work was funded by the BBSRC (grant number BBS/B/01251).

## References

- Allman, M. J., Ward-Robinson, J., & Honey, R. C. (2004). Associative Change in the Representations Acquired During Conditional Discriminations: Further Analysis of the Nature of Conditional Learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 118-128.
- Asratyan, E.A. (1961). The initiation and localisation of cortical inhibition in the conditioned reflex arc. *Annals of the New York Academy of Sciences*, 92, 1141-1159.
- Bonardi, C. (1991). Blocking of occasion setting in feature-positive discriminations. *Quarterly Journal of Experimental Psychology*, 43B, 431-448.
- Bonardi, C. (1996). Transfer of occasion setting: The role of generalization decrement. *Animal Learning and Behavior*, 24, 277-289.
- Bonardi, C. (2007). Occasion setting is specific to the CS-US association. *Learning and Motivation*, 38, 208-228.
- Bonardi, C., Bartle, C., & Jennings, D. (2012). US specificity of occasion setting: Hierarchical or configural learning? *Behavioural Processes*, 90, 311-322.



- Bonardi, C., & Hall, G. (1994a). Occasion-setting training renders stimuli more similar: Acquired equivalence between the targets of feature-positive discriminations. *Quarterly Journal of Experimental Psychology*, *47B*, 63-81.
- Bonardi, C., & Hall, G. (1994b). Discriminative inhibition is specific to the response-reinforcer association but not to the discriminative stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 278-291.
- Bonardi, C., & Jennings, D. (2007). Occasion setting of timing behaviour. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*, 339-348.
- Bonardi, C., & Jennings, D. (2009). Learning about associations: Evidence for a hierarchical account of occasion setting. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 440-445.
- Bonardi, C., & Ward-Robinson, J. (2001). Occasion setters: Specificity to the US and the CS-US association. *Learning and Motivation*, *32*, 349-366.
- Bouton, M.E., & Nelson, J.B. (1998). Mechanisms of feature-positive and feature-negative discrimination learning in an appetitive conditioning paradigm. In N.A. Schmajuk & P.C. Holland (Eds.) *Occasion setting: Associative learning and cognition in animals*. (pp.69-112). Washington D.C.: American Psychological Association.

- Brandon, S.E., Vogel, A.H., & Wagner, A.R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioral Processes*, 110, 67-72.
- de Brugada, I., Garcia-Hoz, V., Bonardi, C., & Hall, G. (1995). The role of stimulus ambiguity in conditional learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 21, 275-284.
- Colwill, R.M., & Rescorla, R.A. (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning and Behavior*, 18, 71-82.
- Colwill, R. M. (1991). Negative discriminative stimuli provide information about the identity of omitted response-contingent outcomes. *Animal Learning & Behavior*, 19, 326-336.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology*, 1, 42-45.
- Davidson, T. L., & Rescorla, R. A. (1986). Transfer of facilitation in the rat. *Animal Learning & Behavior*, 14, 380-386.
- Ganesan, R., Pearce, J.M. (1988). Interactions between conditioned stimuli for food and water in the rat. *Quarterly Journal of Experimental Psychology*, 40B, 229-241.

- Haselgrove, M., Robinson, J., Nelson, A., & Pearce, J.M. (2008). Analysis of an ambiguous-feature discrimination. *Quarterly Journal of Experimental Psychology, 61*, 1710–1725.
- Hall, G., & Honey, R. C. (1989). Contextual effects in conditioning, latent inhibition and habituation: Associative and retrieval functions of contextual cues. *Journal of Experimental Psychology: Animal Behavior Processes, 15*, 232-241.
- Holland, P.C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learning & Motivation, 12*, 1-18.
- Holland, P.C. (1983). Occasion-setting in Pavlovian feature positive discriminations. In M.L. Commons, R.J. Herrnstein, & A.R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (Vol. 4, pp. 183-206). New York: Ballinger.
- Holland, P.C. (1985). The nature of conditioned inhibition in serial and simultaneous feature negative discriminations. *Information processing in animals: Conditioned inhibition* (pp. 267-297). Hillsdale, NJ: Erlbaum.
- Holland, P.C. (1986). Temporal determinants of occasion setting in feature-positive discriminations. *Animal Learning and Behavior, 14*, 111-120.
- Holland, P.C. (1989a). Feature extinction enhances transfer of occasion setting. *Animal Learning and Behavior, 17*, 269-279).
- Holland, P.C. (1989b). Acquisition and transfer of conditional discrimination

performance. *Journal of Experimental Psychology: Animal Behavior Processes*, 15,154-165.

Holland, P.C. (1989c). Transfer of negative occasion setting and conditioned inhibition across conditioned and unconditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 311-328.

Holland, P.C. (1992). Occasion setting in Pavlovian conditioning. In D. Medin (Ed.) *The Psychology of Learning and Motivation*. (Vol. 28 pp.69-125). San Diego: Academic Press.

Honey, R.C., & Hall, G. (1989). Acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes*, 15, 338-346.

Honey, R.C., Hall, G., & Bonardi, C. (1993). Negative priming in associative learning: Evidence from serial conditioning procedures. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 90-97.

Honey, R.C., & Ward-Robinson, J. (2002). Acquired equivalence and distinctiveness of cues: I. Exploring a neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 378-387.

Honey, R.C., Watt, A. (1998). Acquired relational equivalence: Implications for the

nature of associative processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 24, 325–334.

Honey, R.C., Watt, A. (1999). Acquired relational equivalence between contexts and features. *Journal of Experimental Psychology: Animal Behavior Processes*, 25, 324–333.

Jenkins, H.M. (1985). Conditioning inhibition of keypecking in the pigeon. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 327-353). Hillsdale, NJ: Erlbaum.

Konorski, J. (1948). *Conditioned reflexes and neuron organisation*. Cambridge University Press.

Konorski, J. (1967). *Integrative activity of the brain*. Chicago University Press.

Kutlu, M.G., & Schmajuk, N.A. (2012). Solving Pavlov's puzzle: Attentional, associative, and flexible configural mechanisms in classical conditioning. *Learning & Behavior*, 40, 269-291.

Mackintosh, N.J. (1983). *Conditioning and Associative Learning*. Oxford University Press.

- Morell, J.R., & Davidson, T.L. (2002). Transfer across unconditioned stimuli in serial feature discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 83-96.
- Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94, 61-73.
- Pearce, J.M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, 101, 587-607.
- Pearce, J.M., George, D.N., & Aydin., A. (2002). Summation: Further assessment of a configural theory. *Quarterly Journal of Experimental Psychology*, 55B, 61–73.
- Rescorla, R.A. (1972). Evidence for "Unique stimulus" account of configural conditioning. *Journal of Comparative and Physiological Psychology*, 85, 331-338.
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, 72, 77-94.
- Rescorla, R.A. (1985). Conditioned inhibition and facilitation. In R.R. Miller & N.E. Spear (Eds.), *Information processing in animals: Conditioned Inhibition*.(pp. 299-326). Hillsdale, NJ: Erlbaum.

Rescorla, R.A. (1986). Facilitation and excitation. *Journal of Experimental Psychology: Animal Behavior Processes*, 12, 325-332.

Rescorla, R. A. (1991a). Combinations of modulators trained with the same and different target stimuli. *Animal Learning & Behavior*, 19, 355-360.

Rescorla, R. A. (1991b). Transfer of inhibition and facilitation mediated by the original target stimulus. *Animal Learning & Behavior*, 19, 65-70.

Rescorla, R.A., (1999). Summation and overexpectation with qualitatively different outcomes. *Learning and Behavior*, 27, 50-62.

Ross, R.T., & Holland, P.C. (1981). Conditioning of simultaneous and serial feature-positive discriminations. *Animal Learning and Behavior*, 9, 293-303.

Skinner, B.F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.

Swartzentruber, D. (1995). Modulatory mechanisms in Pavlovian conditioning. *Animal Learning and Behavior*, 23, 123-143.

### **Vogel Ponce Wagner Ref**

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in*

*animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.

Ward-Robinson, J., & Hall, G. (1999). The role of mediated conditioning in acquired equivalence. *Quarterly Journal of Experimental Psychology B*, *52B*, 335-350.

Watt, A., Honey, R.C. (1997). Combining CSs associated with the same or different USs. *Quarterly Journal of Experimental Psychology*, *50B*, 350–367.

Williams, D.A., Sagness, K.E., & McPhee, J.E. (1994). Configural and Elemental Strategies in Predictive Learning, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 694-709.



TRAIN	TEST		
	Same	Different	Target
$OS_1: cs_1 \rightarrow US_{food}$ $cs_1 \rightarrow nothing$ $OS_2: cs_2 \rightarrow US_{food}$ $cs_2 \rightarrow nothing$	$OS_1: cs_1$ $OS_2: cs_1$	$OS_1: cs_2$ $OS_2: cs_1$	$cs_1$ $cs_2$
$OS_1: cs_3 \rightarrow US_{food}$ $cs_3 \rightarrow US_{food}$ $OS_2: cs_4 \rightarrow US_{food}$ $cs_4 \rightarrow US_{food}$	$OS_1: cs_3$ $OS_2: cs_4$	$OS_1: cs_4$ $OS_2: cs_3$	$cs_3$ $cs_4$

*Table 1: Design of Bonardi (1996).  $OS_1$  and  $OS_2$  were 10-s presentations of diffuse auditory or visual stimuli;  $cs_1$ ,  $cs_2$ ,  $cs_3$  and  $cs_4$  were 5-s presentations of different keylights, which immediately followed OS presentations on compound trials.  $US_{food}$  (access to grain) occurred on CS termination on reinforced trials.*

<b>TRAIN 1</b>	<b>TRAIN 2</b>	<b>REVALUE</b>	<b>TEST</b>
<i>OS<sub>1</sub>: cs<sub>1</sub>→US<sub>food</sub></i> <i>cs<sub>1</sub>→nothing</i> <i>OS<sub>2</sub>: cs<sub>2</sub>→US<sub>food</sub></i> <i>cs<sub>2</sub>→nothing</i> <i>cs<sub>3</sub>→US<sub>food</sub></i>	<i>OS<sub>1</sub>: cs<sub>1</sub>→US<sub>food</sub></i> <i>cs<sub>1</sub>→nothing</i> <i>OS<sub>2</sub>: cs<sub>2</sub>→US<sub>food</sub></i> <i>cs<sub>2</sub>→nothing</i> <i>cs<sub>3</sub>→nothing</i>	<i>cs<sub>1</sub>→US<sub>food</sub></i>	<i>cs<sub>2</sub>?</i> <i>cs<sub>3</sub>?</i>

*Table 2: Design of Bonardi & Hall, 1994a. OS<sub>1</sub> was a 3-minute presentation of a visual cue, cs<sub>1</sub>, cs<sub>2</sub> and cs<sub>3</sub> were 10-s auditory cues. Food only followed cs<sub>1</sub> and cs<sub>2</sub> when they occurred during OS<sub>1</sub> and OS<sub>1</sub> respectively; unaccompanied cs<sub>1</sub> and cs<sub>2</sub> were nonreinforced. Presentations of cs<sub>3</sub> were always unaccompanied.*

OS TRAIN	TRANSFER TRAIN	TEST		
		Same	Different	Target
<i>OS<sub>1</sub>: cs<sub>1</sub>→US<sub>1</sub></i> <i>cs<sub>1</sub>→nothing</i> <i>OS<sub>1</sub>→nothing</i> <i>OS<sub>2</sub>: cs<sub>c</sub>→US<sub>2</sub></i> <i>cs<sub>2</sub>→nothing</i> <i>OS<sub>2</sub> →nothing</i>	<i>cs<sub>3</sub>→US<sub>1</sub></i> <i>cs<sub>4</sub>→US<sub>2</sub></i>	<i>OS<sub>1</sub>: cs<sub>3</sub></i> <i>OS<sub>2</sub>: cs<sub>4</sub></i>	<i>OS<sub>1</sub>: cs<sub>4</sub></i> <i>OS<sub>2</sub>: cs<sub>3</sub></i>	<i>cs<sub>3</sub></i> <i>cs<sub>4</sub></i>

*Table 3: Design of Bonardi Bartle & Jennings (2012) Experiment 1. OS<sub>1</sub> and OS<sub>2</sub> were 10-s visual stimuli, cs<sub>1</sub>, cs<sub>2</sub>, cs<sub>3</sub> and cs<sub>4</sub> 10-s auditory stimuli, and US<sub>1</sub> and US<sub>2</sub> sucrose or oil. There was a 5-s trace interval between OS and CS presentations on compound trials.*

TRAIN 1	TRAIN 2	TRANSFER TRAIN	TEST	
			Same	Different
<i>OS<sub>1</sub>: cs<sub>1</sub>→nothing</i> <i>OS<sub>1</sub>→nothing</i> <i>cs<sub>1</sub>→US<sub>1</sub></i> <i>OS<sub>2</sub>: cs<sub>2</sub>→nothing</i> <i>OS<sub>2</sub>→nothing</i> <i>s<sub>2</sub> →US<sub>2</sub></i>	<i>OS<sub>1</sub>: cs<sub>1</sub>→nothing</i> <i>OS<sub>1</sub>→US<sub>1</sub></i> <i>cs<sub>1</sub>→US<sub>1</sub></i> <i>OS<sub>2</sub>: cs<sub>2</sub>→nothing</i> <i>OS<sub>2</sub>→US<sub>2</sub></i> <i>cs<sub>2</sub>→US<sub>2</sub></i>	   <i>cs<sub>3</sub>→US<sub>1</sub></i>  <i>cs<sub>4</sub>→US<sub>2</sub></i>	   <i>OS<sub>1</sub>: cs<sub>3</sub></i> <i>OS<sub>1</sub>: cs<sub>4</sub></i> <i>OS<sub>2</sub>: cs<sub>4</sub></i> <i>OS<sub>2</sub>: cs<sub>3</sub></i>	

*Table 4: OS<sub>1</sub> and OS<sub>2</sub> were 10-s visual stimuli, cs<sub>1</sub>, cs<sub>2</sub>, cs<sub>3</sub> and cs<sub>4</sub> 10-s auditory stimuli, and US<sub>1</sub> and US<sub>2</sub> sucrose or groundnut oil. There was a 5-s trace interval between OS and CS presentations on compound trials.*

<b>TRAIN 1</b>	<b>Same</b>	<b>Different</b>	<b>TEST</b>
$OS_A: cs_1 \rightarrow US_1$	$OS_A: cs_1 S \rightarrow US_1$	$OS_A: cs_1 D \rightarrow US_2$	<i>S</i> <i>D</i>
$OS_A: cs_2 \rightarrow US_2$	$OS_A: cs_2 S \rightarrow US_2$	$OS_A: cs_2 D \rightarrow US_2$	
$cs_1 \rightarrow nothing$			
$cs_2 \rightarrow nothing$	$OS_B: cs_1 S \rightarrow US_2$	$OS_A: cs_1 D \rightarrow US_1$	
$OS_B: cs_1 \rightarrow US_2$	$OS_B: cs_2 S \rightarrow US_1$	$OS_A: cs_2 D \rightarrow US_1$	
$OS_B: cs_2 \rightarrow US_1$			

*Table 5: Design of Bonardi & Ward-Robinson (2001).  $OS_A$  and  $OS_A$  were 10-s presentations of diffuse auditory or visual stimuli;  $cs_1$ ,  $cs_2$ ,  $S$  and  $D$  were 5-s presentations of different keylights, and immediately followed  $OS$  presentations on compound trials.  $US_1$  and  $US_2$  were red and white lentils.*

TRAIN	TEST
$cs_1 \rightarrow US_1$ <i>OS: <math>cs_1 \rightarrow nothing</math></i>	<i>Group Same</i> <i>OS: <math>cs_1 \rightarrow US_1</math> <math>cs_1 \rightarrow nothing</math></i> <i>OS: <math>cs_2 \rightarrow US_2</math> <math>cs_2 \rightarrow nothing</math></i>
$cs_2 \rightarrow US_2$ <i>OS: <math>cs_2 \rightarrow nothing</math></i>	<i>Group Different</i> <i>OS: <math>cs_1 \rightarrow US_2</math> <math>cs_2 \rightarrow nothing</math></i> <i>OS: <math>cs_2 \rightarrow US_1</math> <math>cs_1 \rightarrow nothing</math></i>

*Table 6: Design of Bonardi, 2007. OS was a 10-s visual stimuli,  $cs_1$  and  $cs_2$  10-s auditory stimuli, and  $US_1$  and  $US_2$  sucrose or groundnut oil. There was a 5-s trace interval between OS and CS presentations on compound trials.*

<b>TRAIN</b>	<b>REVALUE</b>	<b>OS TEST</b>	<b>ASSOCIATION TEST</b>
<i>OS<sub>1</sub>: cs<sub>1</sub> → US<sub>food</sub></i> <i>OS<sub>1</sub>: cs<sub>2</sub> → nothing</i> <i>OS<sub>3</sub>: cs<sub>1</sub> → US<sub>food</sub></i> <i>OS<sub>3</sub>: cs<sub>2</sub> → nothing</i>  <i>OS<sub>2</sub>: cs<sub>2</sub> → US<sub>food</sub></i> <i>OS<sub>2</sub>: cs<sub>1</sub> → nothing</i>  <i>OS<sub>4</sub>: cs<sub>2</sub> → US<sub>food</sub></i> <i>OS<sub>4</sub>: cs<sub>1</sub> → nothing</i>	   <i>OS<sub>1</sub> → US<sub>shock</sub></i>  <i>OS<sub>2</sub> → nothing</i>	   <i>OS<sub>3</sub> ?</i>  <i>OS<sub>4</sub> ?</i>	   <i>cs<sub>1</sub> → US<sub>food</sub></i> <i>cs<sub>2</sub> → US<sub>food</sub></i>  <i>or</i>  <i>cs<sub>1</sub> → nothing</i>  <i>cs<sub>2</sub> → nothing</i>

*Table 7: Design of Bonardi & Jennings (2009). OS<sub>1</sub>, OS<sub>2</sub>, OS<sub>1</sub> and OS<sub>4</sub> were 10-s auditory cues, cs<sub>1</sub> and cs<sub>2</sub> 10-s visual cues; CS onset coincided with OS offset on training trials.*

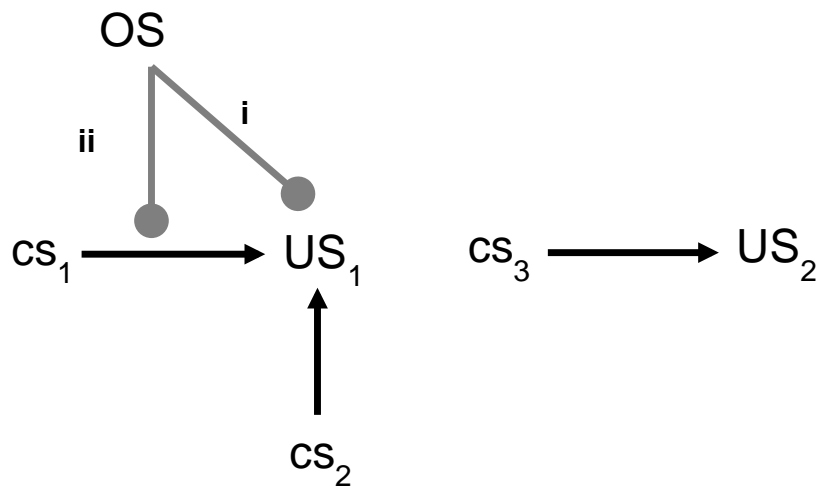
TRAIN 1	TRAIN 2	TRANSFER TRAIN	TEST	
			Same	Different
$OS_1: cs_1 \rightarrow US_1$ $OS_1 \rightarrow nothing$ $cs_1 \rightarrow nothing$	$OS_1: cs_1 \rightarrow US_1$ $OS_1 \rightarrow US_1$ $cs_1 \rightarrow nothing$	$cs_3 \rightarrow US_1$ $cs_4 \rightarrow US_2$	$OS_1: cs_3$ $OS_2: cs_4$	$OS_1: cs_4$ $OS_2: cs_3$
$OS_2: cs_2 \rightarrow US_2$ $OS_2 \rightarrow nothing$ $cs_2 \rightarrow nothing$	$OS_2: cs_2 \rightarrow US_2$ $OS_2 \rightarrow US_2$ $cs_2 \rightarrow nothing$	$POS_1 \rightarrow US_1$ $POS_2 \rightarrow US_2$	$POS_1: cs_1$ $POS_2: cs_2$	$POS_1: cs_2$ $POS_2: cs_1$

Table 8: Bonardi, Bartle & Jennings (2012) Experiments 3a and 3b:  $OS_1$ ,  $OS_2$ ,  $POS_1$  and  $POS_2$  were 10-s visual stimuli,  $cs_1$ ,  $cs_2$ ,  $cs_3$  and  $cs_4$  10-s auditory stimuli, and  $US_1$  and  $US_2$  sucrose or groundnut oil. There was a 5-s trace interval between OS and CS presentations on compound trials.

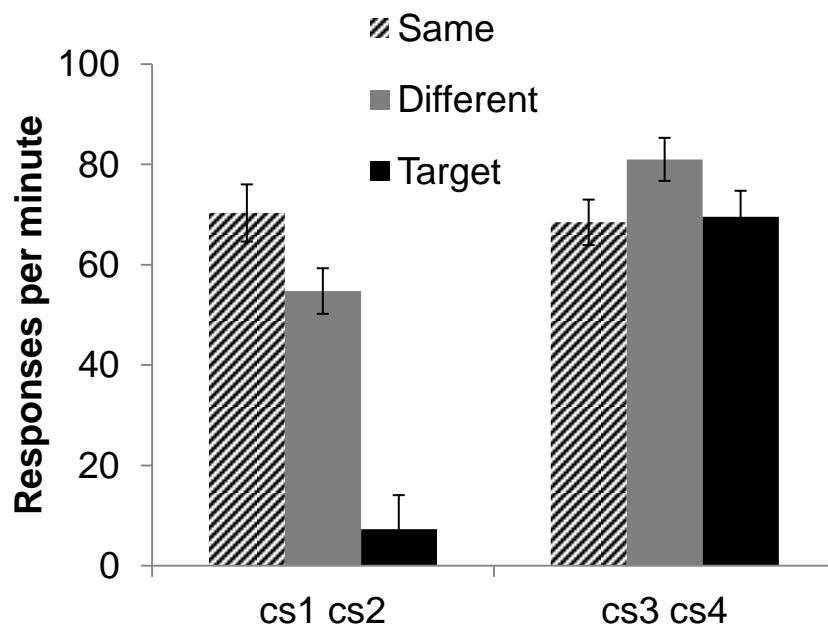


TRAIN 1	TRAIN 2	TEST
$cs_1 \rightarrow US_1$ $cs_2 \rightarrow US_2$  <i>OS: <math>cs_1 \&amp; cs_2 \rightarrow nothing</math></i>	$cs_1 \rightarrow US_2$ $cs_2 \rightarrow US_1$	<p style="text-align: center;"><b>Group Same</b></p> $OS: cs_1 \rightarrow US_1$ $OS: cs_2 \rightarrow US_2$ $cs_1 \rightarrow nothing$ $cs_2 \rightarrow nothing$
		<p style="text-align: center;"><b>Group Different</b></p> $OS: cs_1 \rightarrow US_2$ $OS: cs_2 \rightarrow US_1$ $cs_1 \rightarrow nothing$ $cs_2 \rightarrow nothing$

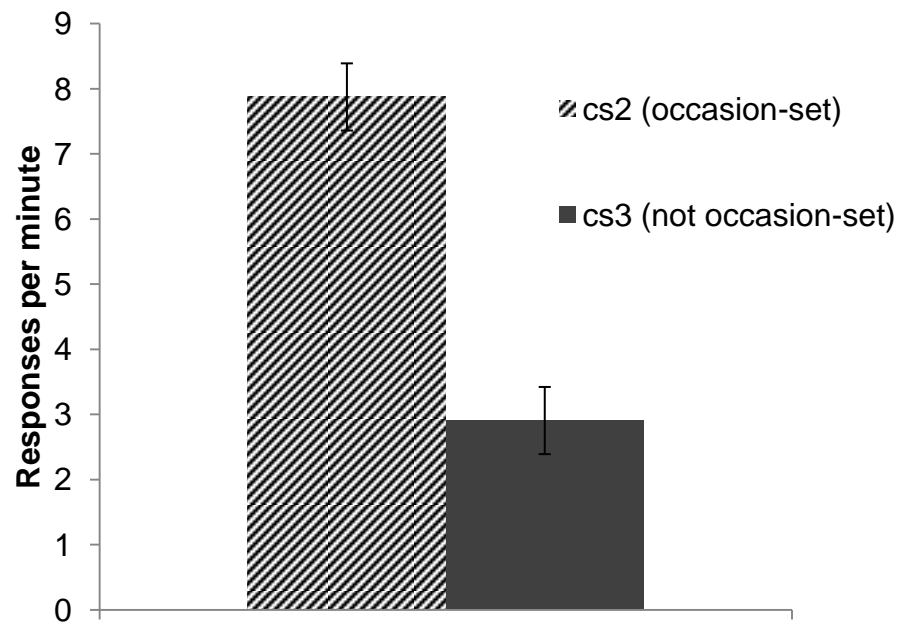
*Table 9: OS was a 10-s visual stimulus,  $cs_1$  and  $cs_2$  10-s auditory or visual stimuli, and  $US_1$  and  $US_2$  sucrose or groundnut oil. There was a 5-s trace interval between OS and CS presentations on compound trials.*



*Figure 1: According to hierarchical accounts an OS signalling that  $cs_1$  will be followed by  $US_1$  acts either on the  $US_1$  representation (i; e.g., Rescorla, 1985) or on the  $cs_1 \rightarrow US_1$  link (ii, e.g., Holland, 1983). The accounts differ in their predictions about the OS's action on a stimulus  $cs_2$  that has signalled  $US_1$ ; neither predicts that the OS will have an effect on a stimulus  $cs_3$  that signals  $US_2$ . The pointed arrow indicates an associative link; round-headed arrow indicates facilitation of (i) activation of  $US_1$  representation or (ii) transmission of activation via  $cs_1 \rightarrow US_1$  association. For further details, see text.*



*Figure 2: Results of Bonardi 1996 Experiment 2: Group mean response rates to the occasion-set  $cs_1$  and  $cs_2$  and the non-occasion-set  $cs_3$  and  $cs_4$  on same ( $OS_1:cs_1$ ,  $OS_2:cs_2$ ,  $OS_1:cs_3$ ,  $OS_2:cs_4$ ) and different ( $OS_1:cs_2$ ,  $OS_2:cs_1$ ,  $OS_1:cs_4$ ,  $OS_2:cs_3$ ) trials or when presented alone. Error bars show within-subject confidence intervals (Cousineau, 2005). For further details, see text and Table 2.*



*Figure 3: Results of Bonardi & Hall, 1994a Experiment 1. Mean response rates to  $cs_2$  (which had been occasion-set) and  $cs_3$  (not occasion-set, but conditioned and extinguished) in the generalisation test. Error bars show within-subject confidence intervals (Cousineau, 2005). For further details, see text and Table 2.*

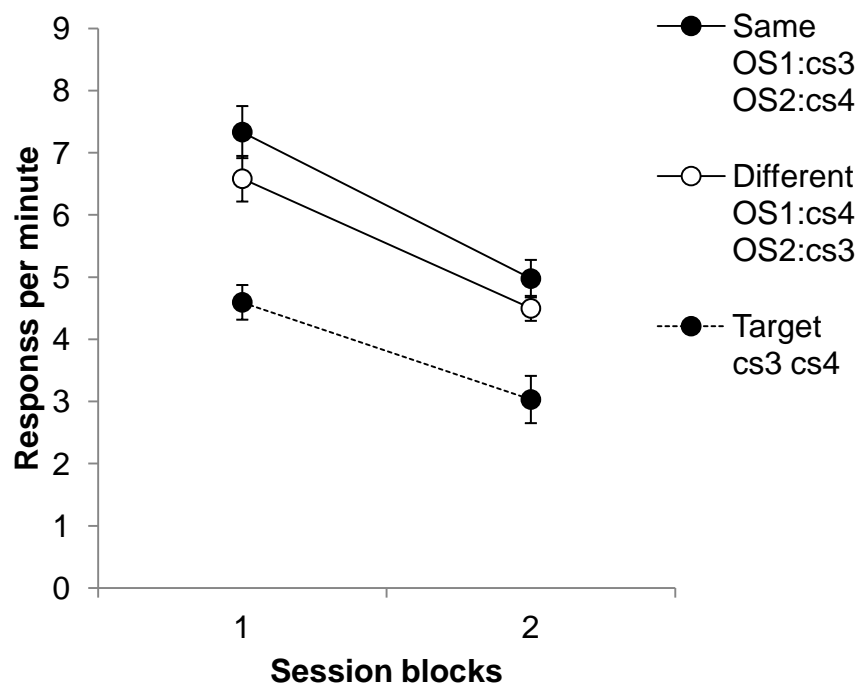


Figure 4: Group mean response rates to  $cs_3$  and  $cs_4$  on same, different and target trials in the two test session blocks of Bonardi, Bartle & Jennings (2012) Experiment 1. Error bars show within-subject confidence intervals (Cousineau, 2005). For further details, see text and Table 3.

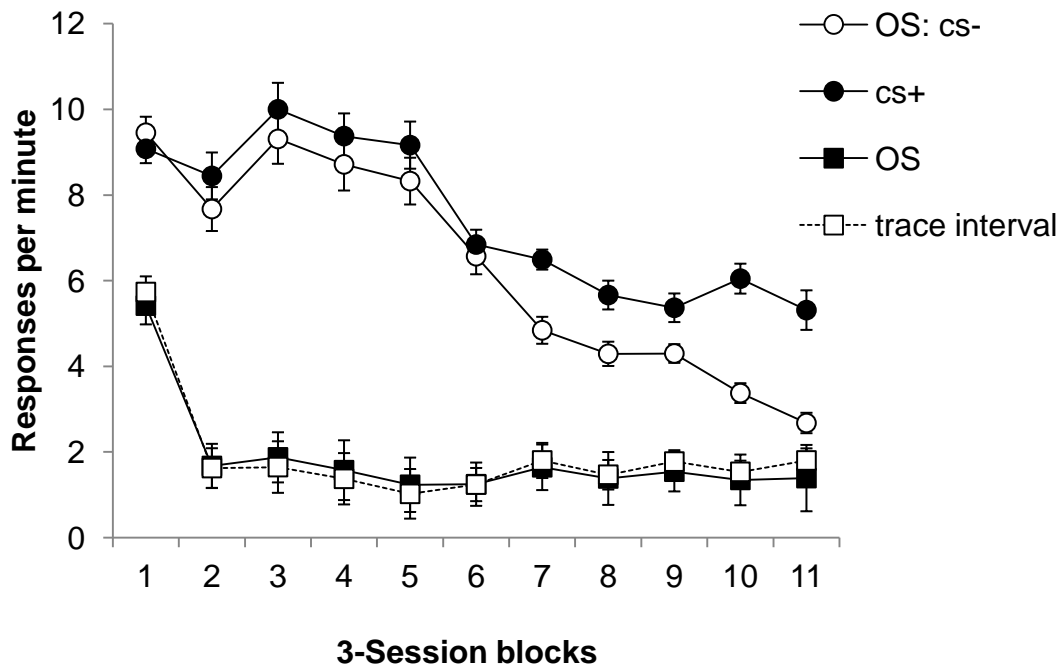


Figure 5: Mean response rates during  $cs_1$  and  $cs_2$  when nonreinforced and signalled by  $OS_1$  and  $OS_2$  ( $OS_1$ :  $cs_1^-$ ,  $OS_2$ :  $cs_2^-$ ) or when presented alone and reinforced ( $cs_1 \rightarrow US_1$ ,  $cs_2 \rightarrow US_2$ ), and also during the occasion setters ( $OS_1$  and  $OS_2$ ) and the trace interval separating OS and each cs on compound trials, during the training stage. Error bars show within-subject confidence intervals (Cousineau, 2005). For further details, see text and Table 4.

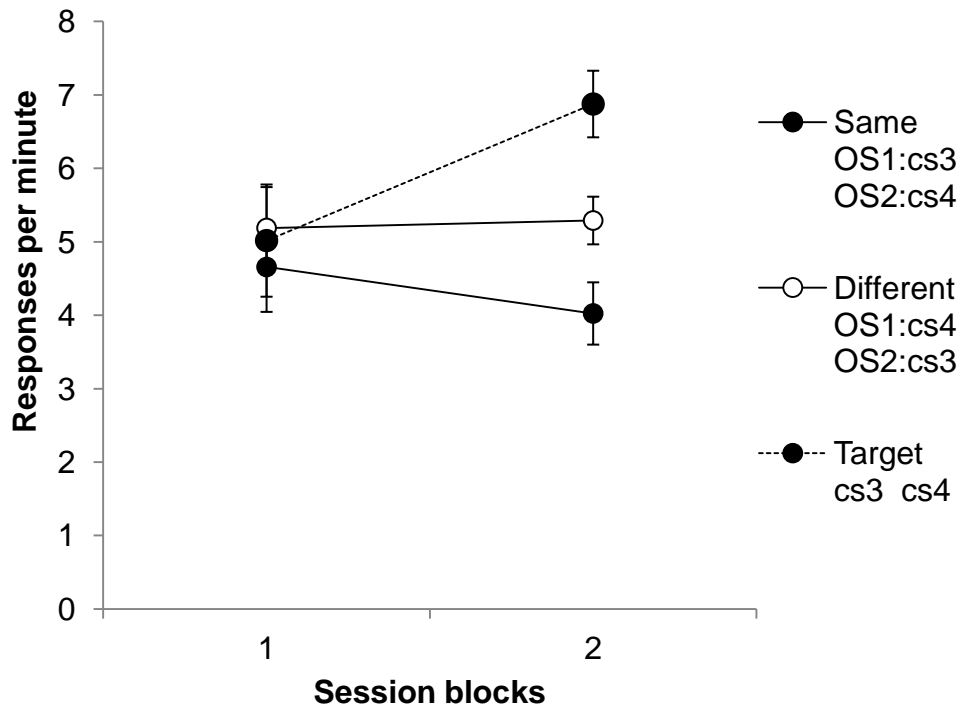
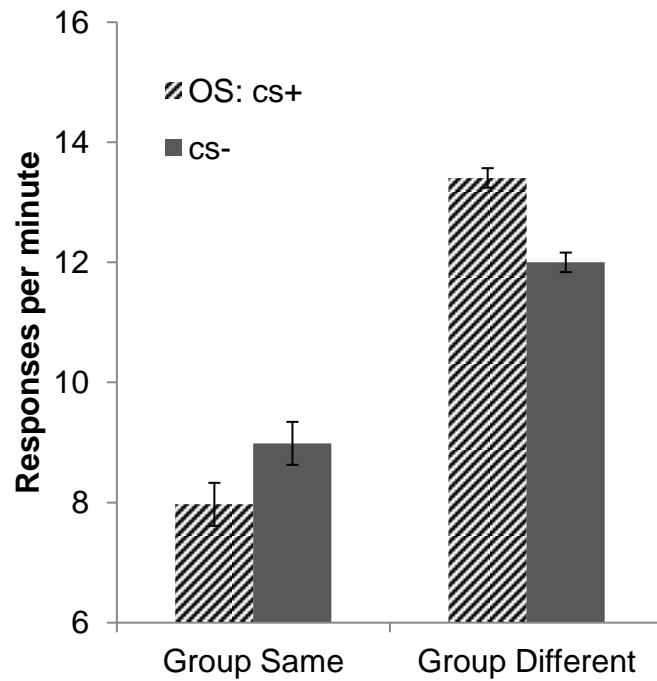
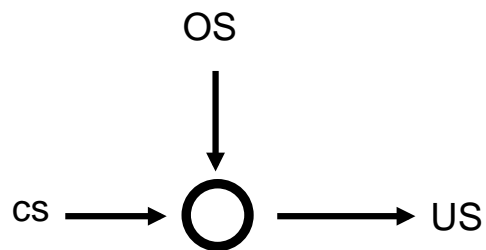
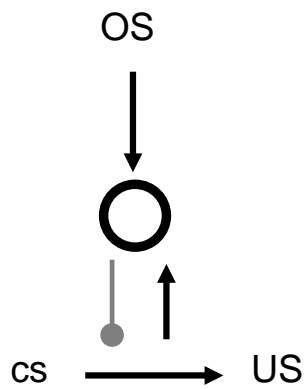


Figure 6: Mean response rates to  $cs_3$  and  $cs_4$  on same, different and target alone trials during the two test session blocks. Error bars show within-subject confidence intervals (Cousineau, 2005). For further details, see text and Table 4.



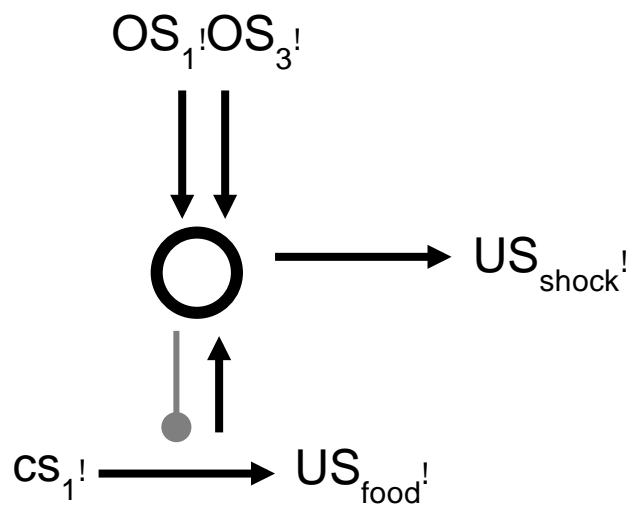
*Figure 7: Group mean rates of responding to  $cs_1$  and  $cs_2$  when signalled by OS (OS:cs) and presented alone (cs) in the test session of Bonardi, 2007 Experiment 2. Error bars show within-subject confidence intervals calculated separately for each group (Cousineau, 2005). For further details, see text and Table 6.*





*Figure 8: Top panel: Associative structure according to the elaborated hierarchical account. Pointed arrows indicate associative links; the round-headed arrow indicates the facilitation of activity transmission through the associative link; the circle denotes a hidden unit. Formation of the association recruits the hidden unit.*

*Bottom panel: Associative structure according to the extended configural account. Pointed arrows indicate associative links; the circle denotes a hidden unit. For further details, see text.*



*Figure 9: Associative structure arising from Honey & Watt's (1998) experiment (Table 7), according to the elaborated hierarchical account. Both  $OS_1$  and  $OS_3$  become associated with the hidden unit of the  $CS_1 \rightarrow US_{food}$  association during initial training. When  $OS_1$  is paired with shock the hidden unit becomes associated with the shock, thus allowing future presentations of both  $OS_3$  and  $CS_1 \rightarrow US_{food}$  to activate the shock representation via the hidden unit. Pointed arrows indicate associative links; the round-headed arrow indicates the facilitation of activity transmission through the associative link; the circle denotes hidden unit.*

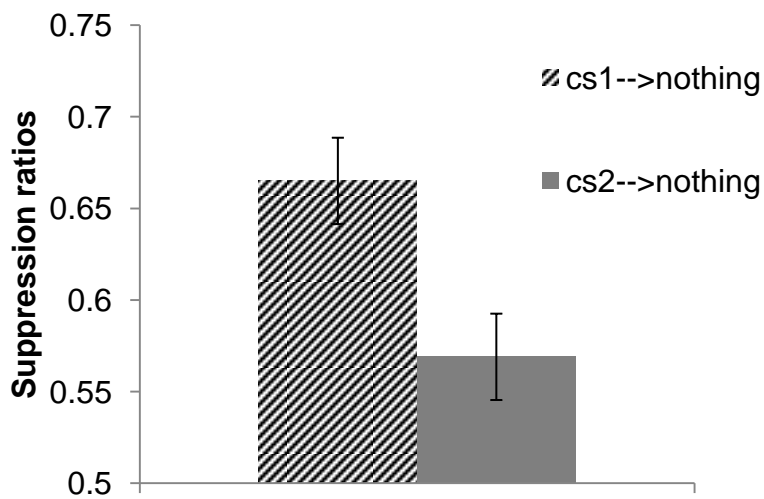
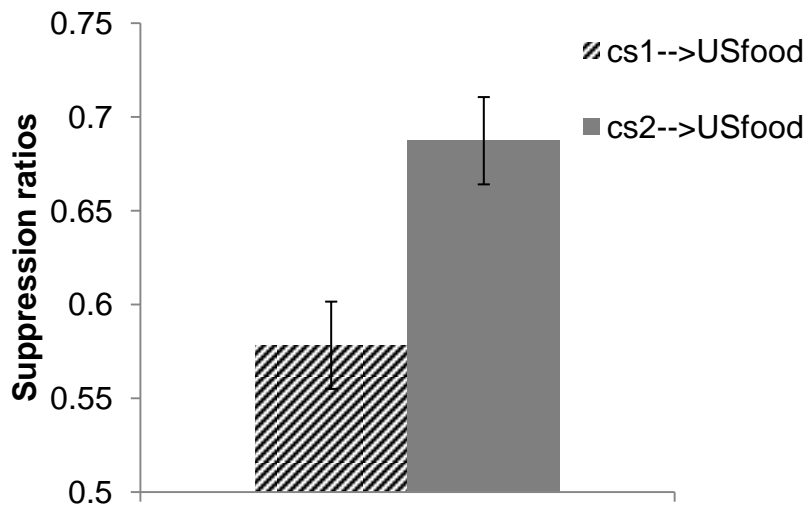
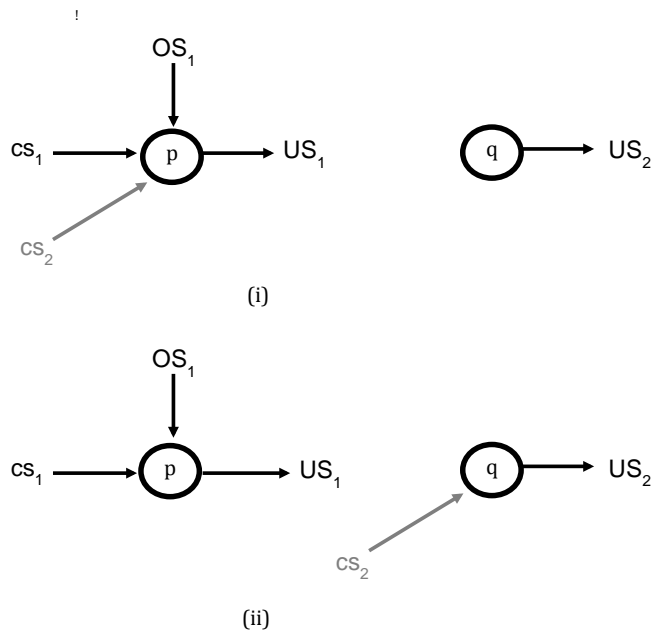
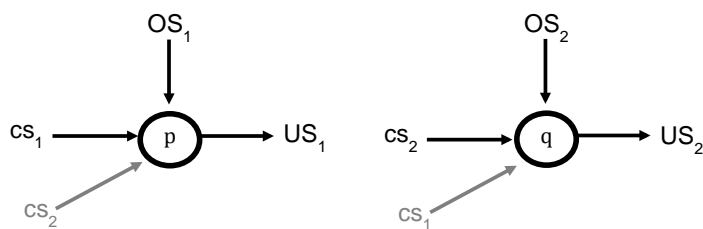


Figure 10: Mean suppression ratios for responding after  $cs_1 \rightarrow US_{food}$  and  $cs_2 \rightarrow US_{food}$  and  $cs_1 \rightarrow nothing$  and  $cs_2 \rightarrow nothing$  in the test sessions of Bonardi & Jennings (2009; Experiments 1 and 2 respectively). Error bars show within-subject confidence intervals (Cousineau, 2005) calculated separately for each group. For further details, see text and Table 7.



*Figure 11a: The extended configural theory account of US specificity. Associative structures arising in feature-positive discriminations in which  $OS_1$  signals reinforcement of  $cs_1$  with  $US_1$ , and (i)  $cs_2$  also predicts  $US_1$  or (ii)  $cs_2$  predicts  $US_2$ . Pointed arrows indicate associative links; circle denotes hidden unit.*



*Figure 11b: Associative structures arising in feature-positive discriminations in which  $OS_1$  signals reinforcement of  $cs_1$  with  $US_1$ ,  $OS_2$  signals reinforcement of  $cs_2$  with  $US_2$ , and then  $cs_1$  is paired with  $US_2$  and  $cs_2$  is paired with  $US_1$ . Pointed arrows indicate associative links; circles denotes hidden units.*

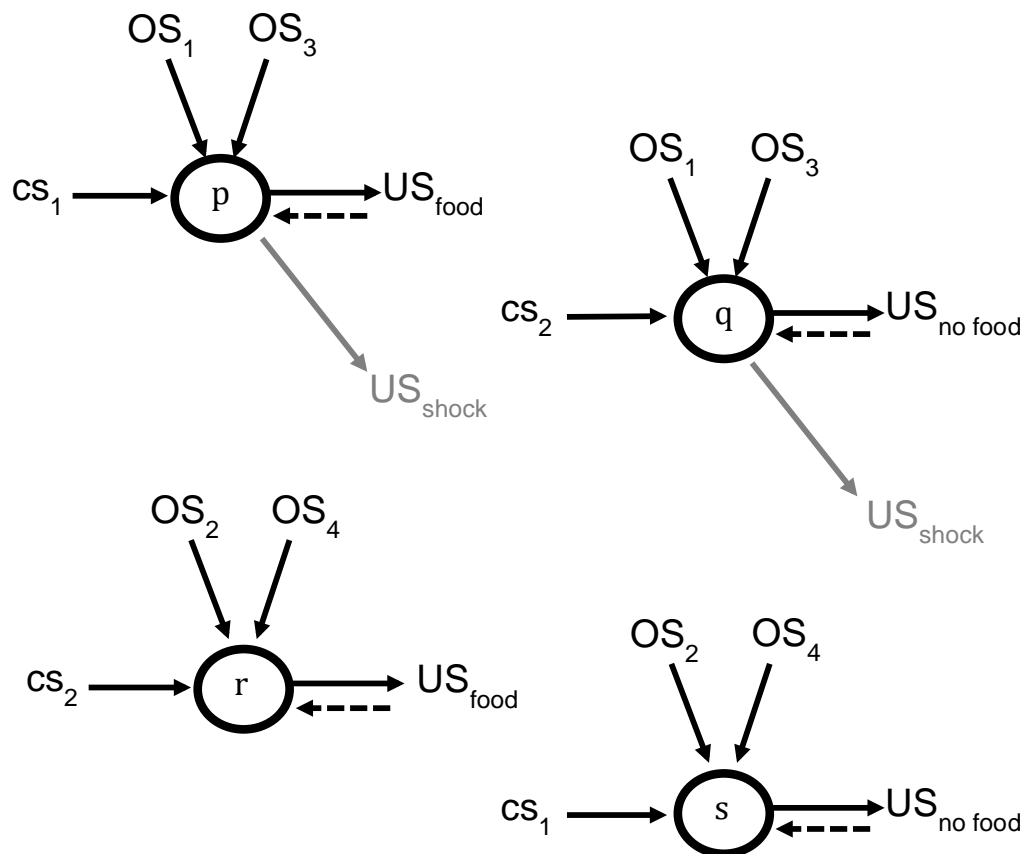


Figure 12: Associative structure arising from Honey & Watt's (1998) experiment (Table 7). According to the extended configural account, both  $OS_1$  and  $OS_3$  signal that  $CS_1$  predicts  $US_{food}$ , so  $OS_1$   $OS_3$  and  $CS_1$  become associated with  $p$ , which is also linked to  $US_{food}$ . Pointed arrows indicate associative links; circles denote hidden units.  $US$  presentation can also activate the hidden unit with which it is linked via a feedback mechanism (dotted arrow).

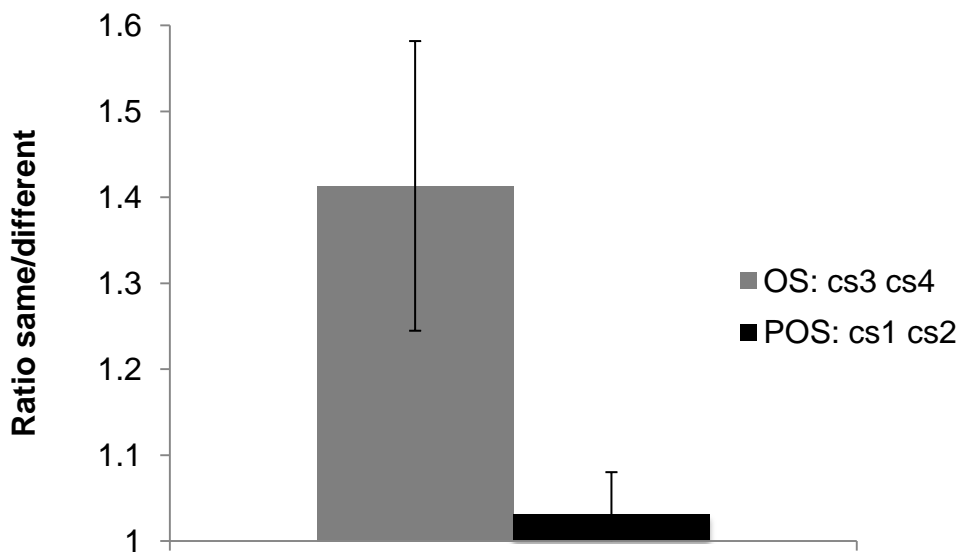


Figure 13: Ratio of same/different responding in Experiments 3a and 3b of Bonardi, Bartle & Jennings, 2012. In Experiment 3a test trials compared responding during transfer exciters  $cs_3$  and  $cs_4$  signalled by the trained occasion setters OS (same  $OS_1$ :  $cs_3$  and  $OS_2$ :  $cs_4$  or different  $OS_1$ :  $cs_4$  and  $OS_2$ :  $cs_3$ ), and in Experiment 3b during the CSs from the occasion-setting discrimination  $cs_1$  and  $cs_2$  signalled by pseudo-occasion setters POS (same  $POS_1$ :  $cs_1$  and  $POS_2$ :  $cs_2$  or different  $POS_1$ :  $cs_2$  and  $POS_2$ :  $cs_1$ ). Error bars show standard error of the mean. For further details, see text and Table 8.

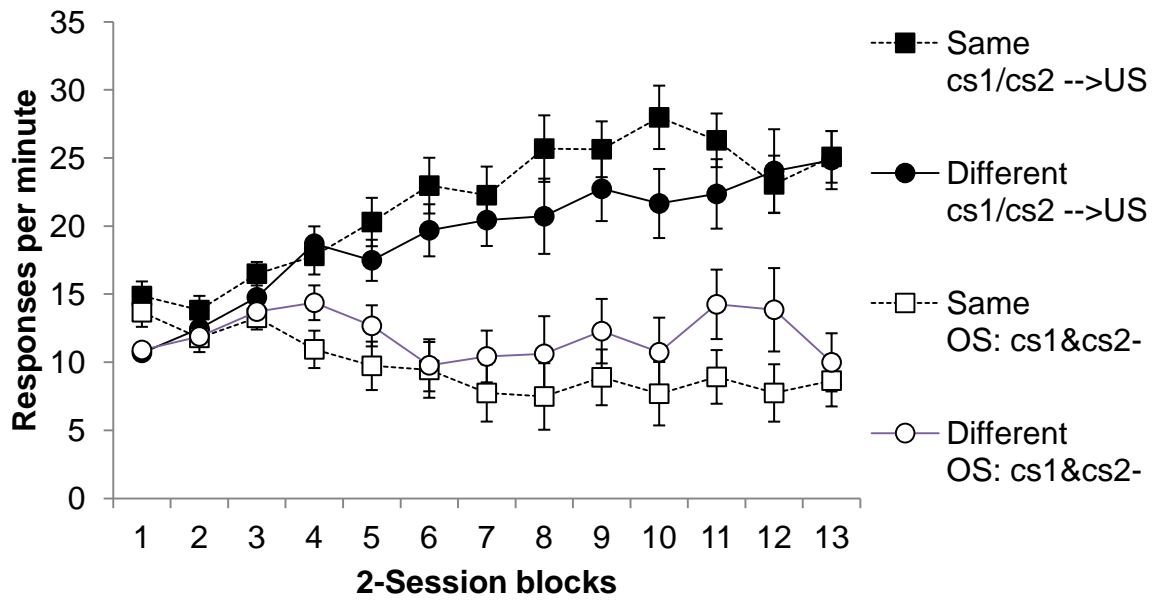
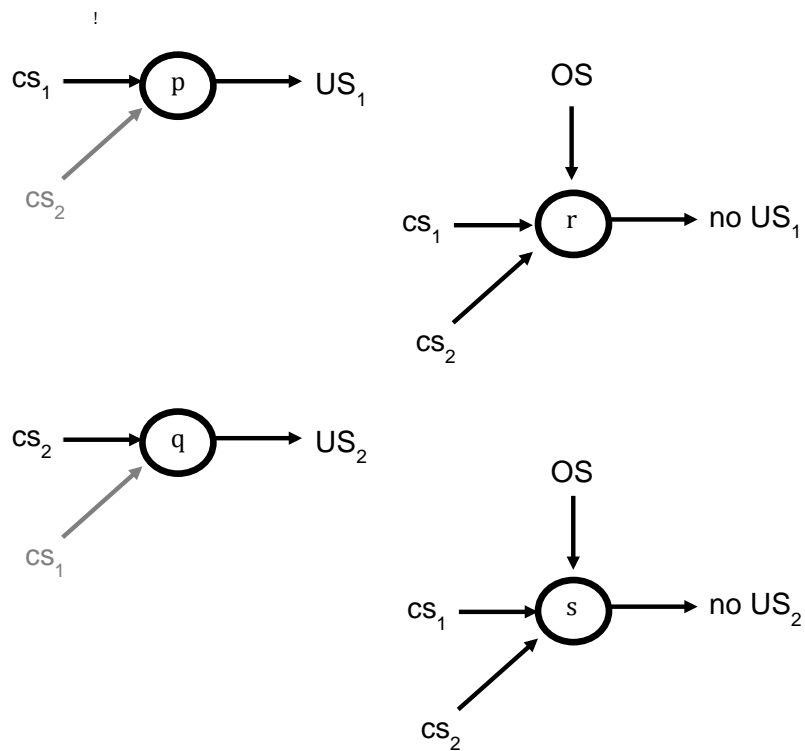


Figure 14: Group mean response rates during  $cs_1$  and  $cs_2$  on reinforced  $cs_1 \rightarrow US_1$   $cs_2 \rightarrow US_2$  and nonreinforced OS:  $cs_1 \& cs_2 \rightarrow \text{nothing}$  trials. Error bars show within-subject confidence intervals (Cousineau, 2005) calculated separately for each group. For further details, see text and Table 9.



*Figure 15: Initial training results in  $cs_1$  and  $cs_2$  becoming linked to  $US_1$  and  $US_2$  respectively by two hidden units,  $p$  and  $q$ . On nonreinforced trials two further hidden units are established, one linking  $OS$ ,  $cs_1$  and  $cs_2$  with  $noUS_1$ , and another linking these same stimuli with  $noUS_2$ . Finally  $cs_1$  and  $cs_2$  are each paired with the alternative outcomes, becoming linked to  $q$  and  $p$  respectively. Pointed arrows indicate associative links; circle denotes hidden unit.*



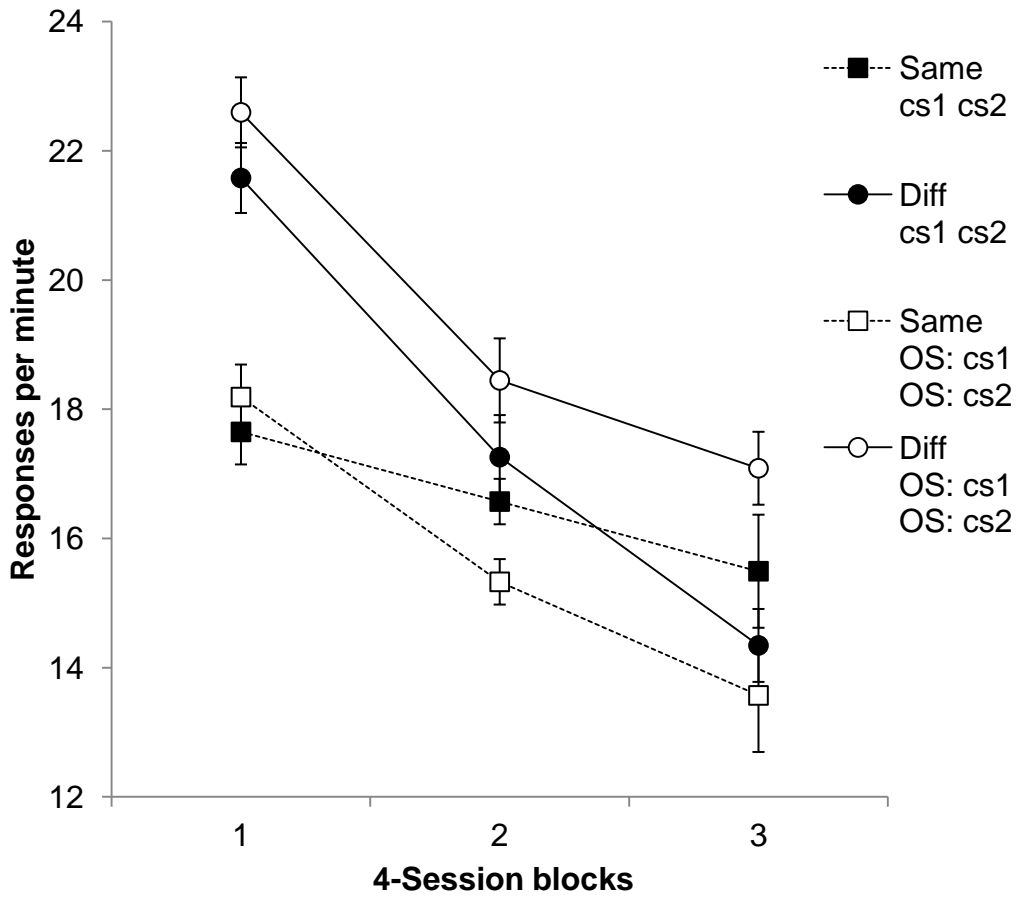


Figure 16: Group mean response rates during the second test phase.  $cs_1$  and  $cs_2$  occurred on reinforced trials, when they were signalled by the occasion setter (OS:  $cs$ ) and on nonreinforced trials, when they were presented alone ( $cs$ ). Error bars show within-subject confidence intervals (Cousineau, 2005) calculated separately for each group. For further details, see text and Table 9.