McLaren, I. P. L., & Verbruggen, F. (in press). Association and inhibition. In R. A. Murphy & R. C. Honey (Eds.), *The Wiley Blackwell handbook on the cognitive neuroscience of learning.* Chichester, England: John Wiley and sons.

Association and Inhibition

Introduction and Manifesto

What is inhibition? The "problem of inhibition" is one that has puzzled learning theorists for many decades. Once it had been demonstrated that pairing a CS (such as a tone or a light) with a US (such as food or shock) produced excitatory conditioning (Pavlov 1927, and see Chapter 2 of Mackintosh 1974), it was natural to consider if a signal could "undo" the effect of an excitatory CS. We now call such a signal a *Conditioned Inhibitor*. A viable recipe for producing conditioned inhibition is to use a design such as A+ AB-, which simply denotes trials where A and the US are paired, interspersed with trials where A and B occur in compound but without the US. The result is that B acquires the properties of being hard to condition to that US (i.e., it passes the retardation test for a conditioned inhibitor), and of suppressing excitatory responding when presented in compound with A or with another excitatory CS that has been conditioned with the same US (i.e., it passes the summation test for conditioned inhibition). In this chapter, we will ask what it is about B that enables it to pass these tests, and what it is about the A+ AB- design that confers these properties. But first we must consider another use of the term "inhibition", one that is just as prevalent amongst cognitive psychologists, but gives a somewhat different meaning to the concept.

Inhibitory control is often invoked in the domain of cognition and action. If one is trying to suppress a thought or withhold an inappropriate or irrelevant action, then we speak of inhibiting that thought or action as part of the solution to the problem. This type of inhibition is considered to be one of the "executive processes" available to us, a deliberate top-down act of control enabling us to cope with ever changing circumstances (e.g., Baddeley, 1996; Logan, 1985; Miyake et al, 2000). As such, the parallel with the research alluded to in the first paragraph, which has often been with rats, rabbits or pigeons as subjects, is not particularly obvious. But more recent research, e.g. Verbruggen and Logan (2008), has found that this act of cognitive control can, in fact, become associatively mediated. In other words, cues that are reliably paired with stopping a response can prime and potentiate that act of control, and may even be able to instigate it in their own right. We shall argue that this is another form of conditioned inhibition, and one of the questions we wish to explore in this chapter is to what extent it shares similarities with the older construct used by learning theorists that goes by the same name.

We begin by reviewing some of the basic properties of conditioned inhibition as studied in animals, and consider the extent to which these phenomena also apply to humans. Our focus then switches to top-down cognitive and motor inhibition and an

evaluation of to what extent it can be associatively mediated. We review the evidence for this phenomenon and again seek to establish some of its basic characteristics. We end by taking an overtly computational perspective on both sets of phenomena as we look for similarities and differences between them.

Basic Phenomena I: Conditioned Inhibition

Conditioning

If we pair an initially neutral stimulus such as a tone or a light (the CS), with a motivationally significant stimulus such as food or shock (the US), then we expect an animal exposed to these contingencies to learn that the CS predicts the US (given that the stimuli are sufficiently salient, the timing between presentation of the CS and US is appropriate etc.). This is demonstrated by means of a change in behaviour of the animal. For example, when the light comes on it may run to the magazine where the food is delivered, or when the tone sounds it freezes, interrupting its current behaviour in preparation for an anticipated shock. These are examples of Pavlovian conditioning; and are conventionally explained by positing that an association from some representation of the CS to some representation of the US has been set up in the animal's mind, such that activation of the CS representation now leads to associatively-mediated activation of the US representation, which is sufficient to generate the observed change in behaviour. This explanation of learning, as being due to the formation of an excitatory link between CS and US representations, is not without its problems, but it does capture many of the basic phenomena of Pavlovian conditioning, including the observation that responses elicited by a trained CS are often similar to that elicited by the US with which it has been paired (cf. Pavlov's principle of stimulus substitution). This principle states that the CS becomes a substitute for the US, and hence elicits a reaction that is similar in its topography to that elicited by presentation of the US itself.

Conditioned Inhibition

Once a CS (denoted as A) has been established as an excitor for a US by means of A+ training (where the + denotes the US), we can use a basic feature-negative design to create a conditioned inhibitor. We simply present the animal with trials in which a compound of A and another CS, namely B, are presented in the absence of the US (AB- trials), whilst still interspersing A+ trials to maintain A as an excitor. B is the "negative feature" in this design, because the otherwise expected reinforcement (predicted by the presence of A) is not delivered when B occurs. One way of expressing this is to say that B has a negative correlation with the US in this design. The consequence of this procedure is that responding to the compound of A and B diminishes over trials and can completely disappear. As a result we infer that B becomes a conditioned inhibitor, able to function as a kind of "safety signal" when the US is aversive (e.g., shock). But initially there was considerable debate about the status of B, because when presented on its own it is quite possible for it to have no detectable effect on behaviour. Indeed, as we shall see, presenting B on its own after this type of training procedure can have little effect on the status of B as well.

Tests for Inhibition

In order to reveal the effects of feature-negative training on B, we conventionally use retardation and summation tests (Rescorla, 1969). Taking the latter test first, this

involves presenting the conditioned inhibitor, B in compound with a quite different CS, C, which is also an excitor for the US. When C is presented on its own it causes the conditioned response associated with that combination of CS and US (e.g., freezing if we are dealing with tone and shock). But if it is presented in compound with B, then this response is diminished, and to a greater extent than if we had simply presented C with D, another CS which is equally familiar but has not been trained as a conditioned inhibitor (or excitor). Thus, we can see that B is able to have an influence over behaviour even in the absence of A, and warrants its status as a conditioned inhibitor in its own right. The retardation test takes a somewhat different approach by pairing B with the US for which it is a conditioned inhibitor. The result is that B+ training proceeds more slowly than D+ training, indicating that some "inhibition" has to be overcome to turn B into an excitor. Thus, both the summation and retardation tests demonstrate that A+ AB- training has changed the status of B from a neutral CS to something that now has an effect which is, in some sense, the opposite to that of an excitor.

Acquisition

One characteristic of conditioned inhibition is that it typically develops more slowly than excitation. Obviously if one has to first establish A as an excitor by means of A+ training before we can use AB- to confer inhibitory properties on B then this necessarily follows for trivial reasons. A more interesting demonstration of this point can be found by comparing acquisition of this feature-negative design with its feature-positive counterpart. Thus, if we contrast the A+ AB- design with C- CD+, in the former B acquires inhibitory control over the discrimination whereas in the latter D develops excitatory control in the feature-positive equivalent. The standard result here is that the feature-positive discrimination is acquired more rapidly than the feature-negative, suggesting that it takes longer to develop B as a conditioned inhibitor than it does D as a conditioned excitor (see Lotz, Uengoer, Koenig, Pearce and Lachnit, 2012).

Another point to note is that it is not necessary to use a full A+ AB- design to make B a conditioned inhibitor; a design of the form A+ AB+ will also work, where A is followed by a greater magnitude of reinforcement (+) than AB (+). The reduction in the reinforcement (or in the probability of reinforcement) is itself enough to confer inhibitory properties on B (Cotton, Goodall and Mackintosh, 1982; Harris, Kwok and Andrew, 2014). These studies, and others like them, suggest that what is crucial in developing conditioned inhibition is that an expectation of one level or rate of reinforcement is contradicted by experience, and that this leads to the development of something quite different to simple excitatory learning. For example, if we were to contrast B in Cotton et al.'s experiment to another stimulus D that had received CD+ training in the absence of any prior training to C, then we would not expect D to have acquired any inhibitory properties (quite the reverse!).

Extinction

Perhaps one of the most eye-catching characteristics of conditioned inhibition is that, according to Rescorla and Zimmer –Hart (1974), inhibitors do not extinguish. After establishing a CS (B) as a conditioned inhibitor, B can be presented on its own for a number of extinction trials, B-, without diminishing its capacity to inhibit (i.e. it will still pass summation and retardation tests). Even if we extend the extinction procedure to a point well beyond that needed to reduce responding to an excitor to floor, the

inhibitory properties of B persist, suggesting once again that there is something rather different about an inhibitory association when contrasted with an excitatory one (which extinguish very readily).

Mediated Inhibition: The Espinet Effect

Inhibition can manifest in conventional CS-US designs as well as in what are in effect simple sensory preconditioning designs. If we pre-expose two sets of compound stimuli (e.g., a solution of sucrose+lemon and another of saline+lemon; AX and BX) then a straightforward analysis of the stimulus contingencies leads to the conclusion that the saline and the sucrose features of these stimuli should come to inhibit one another because of the negative correlation between their presentation: whenever the sucrose (A) occurs, the saline (B) does not, and vice versa (see McLaren, Kaye and Mackintosh, 1989; McLaren and Mackintosh, 2000, 2002; and McLaren Forrest and Mackintosh, 2012, for a more detailed analysis). More specifically, as a result of pairing A and X, X becomes associated with A, and when we now present BX, we have a recipe for establishing an inhibitory association from B to A (because B signals the absence of A). A similar process will establish inhibitory associations from A to B. We can reveal the existence of these mediated inhibitory associations by conditioning A (Espinet, Iraola, Bennett and Mackintosh, 1995). After a few A+ trials (pairing sucrose with lithium chloride to make the animal feel ill) the animal will become averse to drinking A. But when solution B is subsequently tested, then we find no aversion relative to controls. Furthermore, B passes the summation and retardation tests: it reduces aversion to another CS, C, which has also been paired with LiCl, when tested in compound with it (summation test), and is itself harder to condition an aversion to than another flavour, D (retardation test). This is the Espinet effect, and the most plausible interpretation of these results is that B has the ability to depress the activity of A via an inhibitory association with A, and that this then in turn expresses itself via the association between A and the US but with the opposite sign to normal excitatory activation. Thus, what we have in effect here is an example of mediated conditioning, but with the mediation via an inhibitory rather than an excitatory association. Later on we will argue that this result and others like it require a particular implementation of an inhibitory association that differs from that more commonly involved in conditioned inhibition.

The reason we are able to assert this last conclusion is that Bennett, Scahill, Griffiths and Mackintosh (1999) have shown that the effect is asymmetric with respect to which of A or B is conditioned after alternating exposure to AX and BX. If the exposure is such that on each day experience of AX is always followed by BX, but then there is no further trial until the next day, our analysis implies that the inhibitory B->A association should be strong, but that from A->B should be relatively weak. This is because the AX trial leads to a strong X->A association, which allows the development of an inhibitory B->A association, but the B->X association will have decayed considerably before AX is experienced on the next day reducing learning of the inhibitory A->B association. If we now condition A after this pre-exposure to AX and BX, we find good evidence that B has acquired inhibitory properties. Our explanation of this is that the inhibitory link from B->A can activate a representation of A in such a way as to depress the US representation now associated with A. But if instead we were to condition B, we would find little evidence of A acquiring

inhibitory properties, suggesting that the lack of an inhibitory link from A to B prevents the Espinet effect from occurring in this case.

Backward Conditioned Inhibition

One version of the basic conditioned inhibition procedure can be summarised as A+ | AB-. If conditioning A is followed by compound presentations of A with B in the absence of the US, B becomes inhibitory. This design can be more fully characterised as Forward Conditioned Inhibition. Backward Conditioned Inhibition simply involves reversing the ordering of presentation of A+ and AB-, thus AB- | A+. Remarkably, the effect is very similar to that obtained with a forward design, namely that B becomes inhibitory. This effect was discovered in humans by Chapman (1991) and subsequently replicated and further investigated by Le Pelley, Cutler and McLaren (2000). It is not susceptible to the same explanation as that offered for the Espinet effect as the association between A and B in this case must be excitatory. Thus, an explanation in terms of associatively retrieved representations entering into learning with the opposite sign to perceptually activated representations (e.g., modified SOP, Dickinson and Burke, 1996; negative alpha, Van Hamme and Wasserman, 1994), post-acquisition comparison (Miller and Schactman, 1985) or memory-based effects as a consequence of retrieval (Le Pelley and McLaren, 2001), must be deployed. We do not have space here to discuss these alternative explanations of the phenomenon, but simply note that it exists, and that the backward procedure is another effective method for producing inhibitory effects.

Inhibition in Humans

It is worth stating that most of the effects we have considered so far can be demonstrated in humans. For Backward Conditioned Inhibition see Le Pelley and McLaren, 2001; Le Pelley et al, 2000; and also Graham, Jie, Minn, McLaren and Wills, 2011 for demonstrations. Graham (1999) obtained the Espinet effect in humans using a medical diagnosis paradigm and demonstrated the asymmetry found by Bennet et al (1999). Similarly, Mundy, Dwyer and Honey (2006) were able to establish the existence of this asymmetry using procedures that closely paralleled those used by Bennet et al (1999) with rats. Thus, these effects seem to be general and characteristic of associative learning across species.

What is learned during inhibitory conditioning?

There are two main accounts of what is learned during inhibitory conditioning. The first account states that subjects learn an inhibitory association between the CS and the US, which suppresses the US representation (Konorski, 1948). The basic idea here is that an inhibitory association is simply a negative excitatory one. This type of associative structure (shown in the left panel of Figure 1) emerges naturally from the Rescorla-Wagner view of conditioning (Rescorla & Wagner, 1972), and from the idea that inhibition is the consequence of a disconfirmed expectation of an outcome. In essence, the contingencies involved in the A+ AB- training lead to the development of the excitatory connection from the representation of A to the US representation, and the inhibitory connection from the representation of B to that same US representation. Thus, excitation is simply the converse of inhibition and vice versa. The fact that there is little evidence for relatively long-distance inhibitory connections at the neural level is not an immediate argument invalidating this architecture, as we can imagine the inhibitory connection being made up of a long-distance excitatory connection directly to an inhibitory neurone that operates at a local level. By "long-distance" connection,

we simply mean a connection between different (distant) brain regions; whereas a short-distance connection refers to a connection between neurons within the same brain region.

The idea of there being a long-distance excitatory connection to some other neurone that then expresses this connection via a local inhibitory interneuron leads fairly straightforwardly to another possible instantiation of inhibition that depends on the existence of mutual antagonism between different centres. This second account posits that, instead of implementing some (relatively) direct negative link from the representation of the inhibitory CS to the US representation, an excitatory link forms from the representation of the inhibitory CS to a "No-US" centre or representation that then inhibits the US representation (e.g. Konorski, 1967; Le Pelley, 2004; Pearce and Hall, 1980). The key difference between this structure and the earlier one is the use of this "No-US" representation making the inhibition in some sense indirect (see the right panel of Figure 1), and the No-US representation is susceptible to at least two different interpretations. In one (favoured by Konorski) the representation is US specific, and so, in the case where A is trained with food pellets, the No-US representation would be "No food pellets", but in the case where A is trained with sucrose, the No-US representation would be "No sucrose". Another approach to implementing the "No-US" account is to first posit that all conditioning is either appetitive or aversive, and that there are "centres" corresponding to this that mutually inhibit one another (e.g. Dickinson & Dearing, 1979; see also Konorski, 1967). These centres can function as the US and No-US centres, with the aversive acting as the No-US centre for appetitive learning and vice versa. This approach depends more on the interaction of two systems that differ in their motivational significance, and as such has more general implications for behaviour, as we shall see. It does not require an ability to target a No-US representation in a US specific fashion, or that there be a distinct No-US representation for each US representation. For this reason, the appetitive/aversive centres approach seems to us to be a better complement to the more direct implementation of conditioned inhibition shown in the left panel of Figure 1.

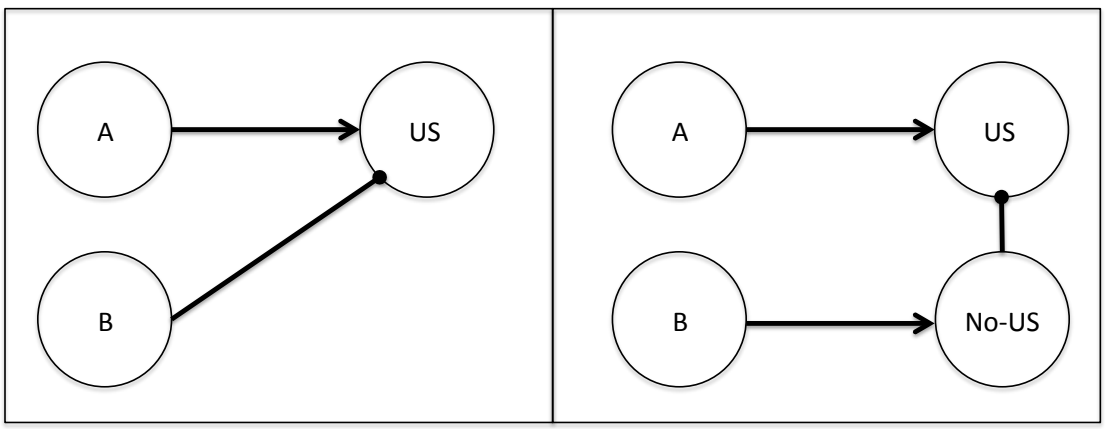

Figure 1: Two different associative structures for the implementation of inhibition. The panel on the left shows a direct inhibitory connection from the representation of the CS to the US representation. The panel on the right shows an indirect inhibitory mechanism whereby the CS representation excites a "No-US" representation that then

inhibits the US representation via an inhibitory interneurone.

We are now in a position to debate these two alternatives, and start by asserting that any account of conditioned inhibition that appeals solely to some interference mechanism is not viable in the light of the evidence available from the animal studies reviewed in this chapter. We can justify this claim by returning to the demonstration by Cotton et al. (1982) showing that conditioned inhibition can be obtained by simply reducing the magnitude of the reinforcer delivered when A and B were presented together (A+ AB+). A tone (playing the role of A) was accompanied by a 1ma shock, and a tone/light compound (AB) was followed by a 0.4 mA shock. The control group either had the tone conditioned alone (followed by a 1 mA shock), or the light conditioned alone (followed by a 0.4 mA shock). This control group is effectively A+ B+. If the apparent inhibition in the experimental group is due to interference caused by the light (B) predicting a 0.4 mA shock rather than a 1 mA shock, then B should produce a similar effect in the B-alone control group. It did not. Clearly there is something special about B in the conditioned inhibition group that stems from the fact that it occurs when a larger shock is expected than that delivered. It's worth noting that the light alone group (A+ B+) in Cotton et al (1982) would quite probably pass the retardation test for inhibition, because we know from Hall and Pearce (1979) that if a tone is first paired with a weak shock then this retards subsequent acquisition of a tone->strong shock relationship. Thus, Cotton et al. have clearly demonstrated that true conditioned inhibition is more than interference. We note that Pearce and Hall (1980) favour an alternative explanation of this result couched in terms of changes in the associability of a stimulus in any case (see also McLaren and Dickinson, 1990).

Additional evidence on this point can be found in the work of Kremer (1978). He showed that compounding a stimulus (B) with stimuli X and Y, which had been separately trained to a given US so that the US was still presented to the compound BXY, conferred inhibitory properties on B. This result relies on the phenomenon of "overexpectation" first demonstrated by Rescorla (1970). If X and Y are both trained individually (X+ Y+) and then trained in compound with the same reinforcer (XY+), the result is that at test X and Y will both elicit less responding in the animal than after the initial training involving the individual stimuli. Thus, a reduction in associative strength is deemed to have taken place as a result of the two stimuli "overpredicting" the US when offered in compound. Kremer predicted that if BXY+ was trained after the X+ Y+ pretraining, then the overexpectation effect should confer inhibitory status on the initially neutral B. Kremer observed exactly this. Our point is that at no stage in this procedure does the outcome (delivery of the same US) change, making any interference account of this phenomenon hard to sustain. This is not to say that interference may not play a role in some demonstrations of what is termed "inhibition", but we do not believe that it can be the full story. This point will take on added significance when we review some of the human data in a later section.

Which brings us back to the question of which of the associative architectures shown in Figure 1 is to be preferred? The evidence that tends to favour the direct link shown in the left-hand panel of Figure 1 is that involving CS-CS associations, such as the Espinet effect. To understand this, it is necessary to realise that the role of A in the figure is being played by the common element X (lemon in this case), the role of B by saline and the role of the US by sucrose. Thus, a pre-exposure trial involving sucrose + lemon leads to an association between their representations forming as shown

between A and the US in the figure. Now a trial following this in which saline + lemon is presented will allow the representation of lemon (A) to activate the representation of sucrose (US), so that the representation of saline (B) forms an inhibitory link to that representation of sucrose (which is not physically present). We have already explained why the effect is thought to be mediated via the ability of saline, say, to inhibit the representation of sucrose after experience of sucrose + lemon/saline + lemon exposure. Clearly, it makes little sense to talk of saline exciting an aversive centre when both the sucrose and saline solutions are essentially neutral prior to conditioning (the rats have a mild liking for both at the concentrations used). We are forced to the conclusion that either the No-US representation has to be very specific (i.e. in this case "No-Sucrose"), or an inhibitory link to the sucrose representation itself is required. Both structures amount to much the same thing once we realise that the "No-Sucrose" structure is effectively an implementation of the direct inhibitory link that gets around the need for relatively long-distance pathways for inhibition (see above). Hence, we are proposing an excitatory link to some local interneuron that then inhibits (locally) the representation of sucrose. Clearly we would also need to postulate some resting activation of this sucrose representation in order for this inhibition mediated via activation of some representation of saline to be effective and to give us the Espinet effect.

The type of evidence that tends to favour the mutually inhibitory appetitive/aversive centres structure draws on studies of trans-reinforcer blocking. Dickinson and Dearing (1979) were able to show that training B to be an inhibitor for a food US enabled it to successfully block learning involving a shock US. That is, once the A+ AB- training was completed using the food US, the next phase was CB+ where the + now denotes shock. Compared to controls, this group learned less about the association between C and shock, suggesting that the prior training of B was, to some extent, blocking acquisition for C. A result of this type fits in well with the idea that the "No-US" centre could indeed be some general appetitive or aversive motivational representation, such that a stimulus that came to predict the absence of food that was otherwise expected could itself acquire aversive properties. It is difficult to see how a result of this type could be generated with the architecture shown in the left-hand panel of Figure 1. For a review of motivational conditioning and interactions between the appetitive and aversive system, see Dickinson and Balleine (2002)

Our final position, then, is that there is evidence for i) a general form of inhibition mediated via excitatory connections to appetitive/aversive centres that mutually inhibit one another and ii) a more specific form of inhibition that is equivalent to a direct inhibitory link to the stimulus representation (be it CS or US) in question. The first mechanism relates more strongly to the motivationally significant stimuli (USs) used in conditioning, the second to structures in what might be termed associative memory.

Basic Phenomena II: Conditioned Inhibitory Control

All our examples of inhibition so far relate to what is called Pavlovian or Classical conditioning where associations are formed between representations of events that occur in the environment. But this is simply one form of what Dickinson calls event-event learning (Dickinson, 1980). Now we turn to the issue of inhibition in an instrumental context, where the task is to withhold or cancel a thought or action rather

than detect the unexpected absence of an event. To do this, we will focus on human experiments that investigate the role of inhibition in executive control. Our review of this area will conclude that in many cases it is unnecessary to appeal to inhibition to explain performance, But there are some circumstances where the case for inhibition seems to be strong, and we will focus on these once we have identified them.

In the last few decades, 'inhibition' has become a central concept in many theories of attentional and executive control. The general tenet is that humans need inhibitory mechanisms to suppress irrelevant stimuli, thoughts, actions, and emotions to effectively deal with the constant inflow of information and multitude of response options. Within the executive control domain, inhibition is not regarded as a unitary construct, and several taxonomies have been proposed. Nigg (2000) distinguished between (1) *cognitive inhibition*, which refers to the suppression of irrelevant thoughts and information in working memory; (2) *interference control*, which refers to suppression of irrelevant stimuli; (3) *behavioural* or *motor inhibition*, which refers to the suppression of automatic, prepared, or cued responses; and (4) *oculomotor inhibition*, which refers to the effortful suppression of reflexive saccades. Similar taxonomies and distinctions between cognitive and behavioural (or motor) inhibition have been proposed by Friedman & Miyake (2004) and Harnishfegher (1995), among others. The case for cognitive inhibition is weak (see e.g. MacLeod et al. , 2003; Raaijmakers & Jakab, 2013). Therefore we will focus on the inhibition of responses.

Top-down response inhibition in interference tasks

The role of inhibition in interference control or congruency tasks, such as the Eriksen flanker task or the Stroop task, is still disputed. Popular dual-route models (e.g., Kornblum, Hasbroucq, & Osman, 1990) assume that responses in congruency tasks are activated via a direct activation route and an indirect activation route. Activation via the direct route is unconditional and automatic, independent of the task instructions. By contrast, activation of the response via the indirect route is deliberate and controlled. Inhibitory accounts state that conflict or interference is resolved by strengthening the processing of relevant information via the indirect route and by selectively inhibiting irrelevant information and responses that were activated via the direct route (e.g., Ridderinkhof, 2002). Some have argued that inhibition is required to suppress all motor responses globally when conflict between alternative actions is detected (Frank, 2006; Wiecki & Frank, 2013). This would effectively allow the system to prevent premature responses and to select the appropriate response.

In recent years, evidence both in favour and against inhibitory accounts of interference control has been forthcoming. First, several studies have demonstrated that top-down inhibition may not be required to resolve interference as this can be achieved by top-down enhancement of relevant information alone. Several computational models of interference control assume that task demand units or representations of the relevant categories will bias processing in the subordinate pathways, enhancing the processing of task-relevant information (e.g., Cohen, Dunbar, & McClelland, 1990; Herd, Banich, & O'Reilly, 2006). It may be that activation of task-relevant information leads to inhibition of competing task-irrelevant processing via lateral inhibitory connections. But it is important to stress that this inhibition is achieved locally and not via top-down inhibitory connections. Thus, inhibition of task-irrelevant information would be a local 'side-effect' of top-down excitation of task-relevant information. Again, this would help to get around the need for relatively

long-distance pathways for inhibition.

But the top-down response-inhibition account has also received support, primarily from neuroscience studies (but see also e.g. Ridderinkhof, 2002). For example, a recent study tested the response inhibition account using motor-evoked potentials (MEPs) elicited by transcranial magnetic stimulation (TMS) of the right motor cortex (Klein, Petitjean, Olivier, Duque, 2014). The authors found reduced MEPs for trials on which the distractors were mapped onto a left response. This suggests that suppression of motor excitability is a component of interference control (see also e.g., van den Wildenberg et al., 2010). It is possible that interference and competition caused by irrelevant stimuli is resolved by activating relevant features and stimulus processing, whereas response competition is resolved by activating the relevant response and selectively suppressing the irrelevant response via separate Go and NoGo pathways between prefrontal cortex and the basal ganglia (e.g., Frank, 2005). More specifically, the relevant response can be activated via activation of 'Go' cells in the striatum, which inhibit the *internal* segment of the globus pallidus (GPi); this reduces inhibition of the thalamus, leading to the execution of a motor response (the direct cortical-subcortical pathway; Nambu et al, 2002) [1]. Irrelevant responses can be suppressed via activation of 'Nogo' striatal cells, which inhibit the *external* segment of the globus pallidus (GPe); this reduces tonic inhibition between GPe and the GPi, resulting in increased activity in GPi, and consequently, increased inhibition of the thalamus (the indirect cortical-subcortical pathway; Nambu et al, 2002). Note that global suppression of all motor output, as postulated by Frank and colleagues, could be achieved via a third pathway, namely the hyperdirect pathway. This involves activation of the subthalamic nucleus, which has in turn a broad effect on GPi, leading to global suppression of the thalamus. Prefrontal areas, such as the presupplementary motor area and the right inferior frontal gyrus, are thought to activate the Nogo cells in the striatum or the subthalamic nucleus.

Aftereffects of top-down inhibition: Negative priming

After a stimulus has appeared as a distractor in congruency tasks such as a picture-naming task or an Eriksen flanker task, responding to it on the next trial is usually impaired. This finding is referred to as 'negative priming'. The dominant inhibition account of negative priming assumes that when an item is a distractor, its representation or the process linking the representation with the response becomes suppressed, and that residual inhibition impairs responding to the item on the following trial (e.g. Tipper, 2001). However, this impairment could be caused by the retrieval of stimulus- and response-information from the previous trial (e.g. Neill, Valdes, Terry, & Gorfein, 1992; Rothermund, Wentura, & De Houwer, 2005). For example, Neill and colleagues proposed that a distractor becomes associated with a do-not-respond representation; when it is repeated on the next trial as a target, the do-not-respond association is activated via associative retrieval, and this will interfere with responding. By contrast, Rothermund et al. (2005) suggested that the distractor becomes associated with the response to the target on the prime trial; retrieval of this response association will interfere with responding on the current probe trial because the retrieved information is usually inconsistent with the currently relevant response (see Jones, Wills and McLaren, 1998, for an example of how this type of response

[1] Note that the cortico-basal-ganglia pathways do not directly map on to the direct and indirect routes discussed in dual-route frameworks.

interference might be implemented). Mayr and Buchner (2007) reviewed the negative priming literature, and argued that the available data generally favour the memory account over the distractor-inhibition account.

There is a parallel to draw between the memory retrieval accounts of negative priming and the conditioned inhibition accounts discussed in Part I. The response-interference account of negative priming is akin to the interference account of conditioned inhibition that assumes US-US interference. As discussed above, interference between CS or US representations may contribute to conditioned inhibition but it seems unlikely that it is the only mechanism responsible for the effects we have covered. Similarly, Rothermund et al., (2005, p.493) noted that 'stimulus-response retrieval is not the only mechanism that produces negative priming, it is one of the underlying mechanisms'. One of the other mechanisms could be the establishment of a link between the stimulus and a 'do not respond' or 'no response' representation, similar to a 'no-US' representation in conditioned inhibition paradigms. This 'no-response' representation could be specific (e.g. 'no left response', akin to a 'no-A' representation) or more general. Consistent with the latter option, Frings, Moeller and Rothermund (2013) have argued that both stimuli and responses may be represented by abstract conceptual codes; for example, responses would be coded in terms of approach or avoidance. In the context of negative priming, this would imply that distractors are linked to a general 'avoid/aversive' representation. Indeed, several recent studies suggest that conflict is aversive (e.g. Fritz & Dreisbach, 2013; van Steenbergen, Band, Hommel, 2009; see also Botvinick, 2007). Furthermore, work by Raymond and colleagues suggest that ignoring a distractor could lead to its' devaluation (e.g. Raymond, Fenske, & Tavassoli, 2003). Again, this is consistent with the idea that stimuli can be linked with general appetitive/approach and aversive/avoidance centres, which mutually inhibit each other. Later on, we will argue that there is good reason to suppose the existence of both mutually inhibitory appetitive/aversive centres and separate approach/avoidance centres, which we will refer to as "go" and "stop" centres.

Top-down inhibition of behaviour

The idea that responses or motor actions can be inhibited in a top-down fashion receives the strongest support from paradigms such as the go/no-go paradigm and the stop-signal paradigm. Therefore, we will focus on these two paradigms in the remainder of this chapter. In the go/no-go paradigm, subjects are presented with a series of stimuli and are told to respond when a go stimulus is presented and to withhold their response when a no-go stimulus is presented (e.g., press the response key for a square but do not press the response key for a diamond; Figure 2, left panel). One could argue that the go/no-go task corresponds to an AX+ | BX- design, with A and B as the go stimulus and the no-go stimulus, respectively, and X as the task context. In the stop-signal paradigm, subjects usually perform a choice reaction task on no-signal trials (e.g., press the left response key for a square and press the right response key for a diamond; Figure 2, right panel). On a random selection of the trials (stop-signal trials), a stop signal (e.g. an auditory tone or a visual cue, such as the outline of the go stimulus turning bold) is presented after a variable delay (stop-signal delay; SSD), which instructs subjects to withhold the response to the go stimulus on those trials. This corresponds to an A+ | AB- design, with A corresponding to the go stimuli, and B the stop signal.

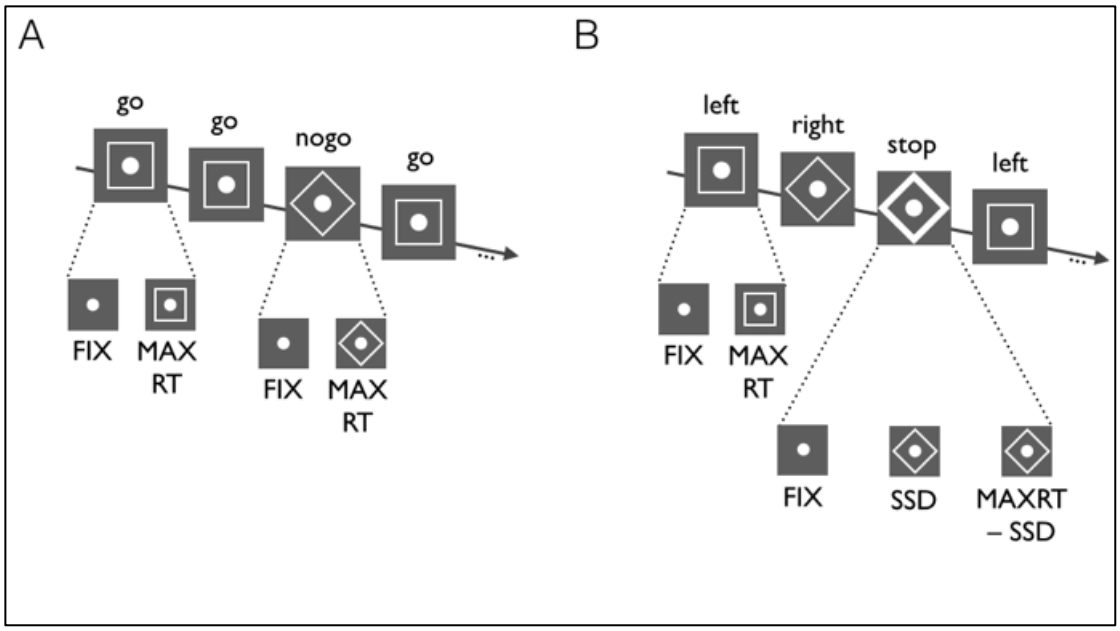

Figure 2: A schematic illustration of the go/no-go and stop-signal paradigms. FIX = duration of the fixation interval; MAX RT = maximum response latency; SSD = variable stop-signal delay in the stop-signal paradigm.

Behaviourally, performance in both paradigms can be modelled as an independent race between a go process, which is triggered by the presentation of a go stimulus, and a stop process, which is triggered by the presentation of the no-go stimulus or the stop signal (Logan & Cowan, 1984; Logan, Van Zandt, Verbruggen, Wagenmakers, 2014; Verbruggen & Logan, 2009). When the stop process finishes before the go process, response inhibition is successful and no response is emitted (signal-inhibit); when the go process finishes before the stop process, response inhibition is unsuccessful and the response is incorrectly emitted (signal-respond). In the stop-signal task, the covert latency of the stop process (stop-signal reaction time or SSRT) can be estimated from the independent race model (Logan & Cowan, 1984). SSRT has proven to be an important measure of the cognitive control processes that are involved in stopping. For recent reviews of studies of response inhibition in cognitive psychology, cognitive neuroscience, developmental science and psychopathology, see e.g. Bari & Robbins (2013b), Chambers, Garavan, & Bellgrove (2009), and Verbruggen & Logan (2008c).

Neurally, response inhibition processes primarily engage a fronto-basal-ganglia inhibition network, which includes the right (and possibly left) inferior frontal gyrus, the pre-supplementary motor area, the anterior cingulate cortex, the dorsolateral prefrontal cortex, parietal regions, and basal ganglia (Aron, Robbins, & Poldrack, 2014; Bari & Robbins, 2013a; Chambers et al., 2009; Swick, Ashley, & Turken, 2011)[2]. On go trials, activation in frontal and parietal areas could lead to activation of a go response via the direct fronto-basal ganglia pathway (see above). In the case of response inhibition, activation in prefrontal areas could lead to a suppression of motor

[2] Inhibition of eye movements may recruit a different network. Single-cell studies indicate that it relies primarily on the activation of movement- and fixation-related neurons in frontal eye fields in dorsolateral prefrontal cortex and superior colliculus in midbrain (for a review, see Schall & Godlove, 2012).

output via the hyperdirect fronto-basal ganglia pathway (see above), resulting in fast and global suppression of motor output. This might affect all response tendencies including activation in muscles that are irrelevant to the task (Badry et al., 2009; Greenhouse, Oldenkamp, & Aron, 2011; Majid, Cai, George, Verbruggen, & Aron, 2012). More selective inhibition of a specific response could potentially be achieved via activation of the indirect fronto-basal pathway (Majid et al., 2012; Smittenaar, Guitart-Masip, Lutti, & Dolan, 2013). The exact cognitive role of the frontal regions is debated, partly because a detailed processing framework is lacking in many neuroscience studies (Verbruggen, McLaren, Chambers, 2014). Moreover, the prefrontal areas that are involved in top-down response inhibition are generally recruited by tasks that require selection of competing actions (Bunge, 2004; Duncan & Owen, 2000) and reprogramming or updating actions (Buch, Mars, Boorman, & Rushworth, 2010; Verbruggen, Aron, Stevens, & Chambers, 2010). Thus, response selection and response inhibition may be two sides of the same coin (see also Mostofsky & Simmonds, 2008), relying on overlapping prefrontal brain areas which bias processing in subordinate systems in a context-dependent fashion.

The independent race model of Logan and Cowan (1984) assumes stochastic independence between the go and stop processes. However, the cognitive neuroscience of stopping indicates that go and stop processes interact to produce controlled movements (see also the discussion of the basal ganglia pathways above). To address this 'paradox', Boucher, Palmeri, Logan, & Schall (2007) proposed an interactive model. In their model, the go process is initiated by the go stimulus and a go representation is activated after an afferent delay. The stop process is initiated by the stop signal and a stop representation is also activated after an afferent delay. Once the stop representation is activated, it inhibits go processing strongly and quickly. In this interactive model, SSRT primarily reflects the period before the stop unit is activated, during which stop and go processing are independent, so its predictions correspond to those of the independent model (Logan & Cowan, 1984).

Conditioned inhibitory control?

Performance in response-inhibition paradigms is usually attributed to a top-down act of control (Verbruggen & Logan, 2008; Verbruggen, McLaren, & Chambers, 2014). However, in recent years, several studies have examined both the short-term and long-term aftereffects of stopping a response. This work suggests that stop representations may be activated via the retrieval of stimulus-stop associations. Eventually, this could lead to automaticity of stopping (Logan, 1988; Verbruggen & Logan, 2008). In other words, inhibitory control may become conditioned.

Several stop-signal studies have observed that response latencies on no-signal trials increase after both successful and unsuccessful stopping. This response slowing has been attributed to strategic control adjustments: subjects must try to find a balance between responding quickly on no-signal trials (speed) and stopping on stop-signal trials (caution); this balance would be adjusted in favour of caution after a stop-signal trial (Bissett & Logan, 2011). However, the slowing is more pronounced when the stimulus or stimulus category of the previous trial is repeated (Bissett & Logan, 2011; Enticott, Bradshaw, Bellgrove, Upton, & Ogloff, 2009; Oldenburg, Roger, Assecondi, Verbruggen, & Fias, 2012; Rieger & Gauggel, 1999; Verbruggen & Logan, 2008a; Verbruggen, Logan, Liefooghe, & Vandierendonck, 2008). This analysis suggests some contribution of memory retrieval. Logan (1988) argued that every time people

respond to a stimulus, processing episodes are stored as instances in memory. These episodes consist of the stimulus (e.g. a shape), the interpretation given to a stimulus (e.g. 'square'), the task goal ('shape judgment'), and the response ('left'). When the stimulus is repeated, previous processing episodes are retrieved, facilitating performance if the retrieved information is consistent with the currently relevant information but impairing performance if the retrieved information is inconsistent. On a stop-signal trial, the go stimulus or stimulus category becomes associated with stopping; when the stimulus (or category) is repeated, the stimulus-stop association is retrieved, and this interferes with responding on no-signal trials. The idea here, then, is that the go response/goal and the stop response/goal are mutually inhibitory (cf. Boucher et al, 2007) in much the way that Dickinson and Dearing (1979) postulate appetitive and aversive stimuli are. This stimulus-stop association account is related to the 'do-not-respond tag' account of the negative priming effect, mentioned earlier (Neill & Valdes, 1992; Neill et al., 1992); of course this is no coincidence because both accounts are based on the Instance Theory of Logan (1988). The stimulus-stop effects are observed up to 20 trials after the presentation of the stop signal (Verbruggen and Logan, 2008a). Similar long-term effects have been observed in task-switching studies, suggesting that stimuli can become associated with tasks or task goals (Waszak, Hommel, & Allport, 2003, 2004, 2005).

Theoretically, repetition priming effects can be viewed as the first step towards automatization (Logan, 1990). According to Logan, automatization involves a transition from performance based on cognitive algorithms or rules to performance based on memory retrieval. Therefore, the observation that a stimulus could prime stopping after a signal trial raises the question whether inhibitory control may become a bottom-up act of control, driven by retrieval of stimulus-stop associations from memory, instead of a top-down act of control. In a series of experiments, we examined the bottom-up idea (Verbruggen & Logan, 2008b). Initially, we used go/no-go tasks in which the stimulus category defined whether subjects had to respond (e.g. natural objects = go) or not (e.g. man-made objects = no-go). We trained subjects to stop their response to a specific stimulus, and then reversed the go/no-go mappings in a test phase. In this test phase, subjects were slower to respond to that stimulus compared with stimuli that they had not seen before (Verbruggen & Logan, 2008b, Experiment 1). Furthermore, learning the new go association was slowed, so one could argue that it passes a retardation test for inhibition. The response slowing was still observed when the tasks changed from training to test: subjects made natural/man-made judgements in training but large/small judgments in test (or vice versa; Experiment 2), and RTs were longer for inconsistent items (i.e. nogo in one task but go in the other task) than for consistent items (i.e. go in both tasks). This last is a result akin to that obtained in summation tests for inhibition if training for a given stimulus in one category was natural+stimulus=nogo, then on test small+stimulus=go; the inhibition derived from training has transferred to the novel test situation in a manner analogous to combining an inhibitor with a novel excitor. We also demonstrated (Experiment 3) that the effect was not entirely category-driven as stimulus-specific slowing was observed when the category-stop mappings were inconsistent in training: here the go/no-go mappings changed every block (e.g. natural=go and man-made = no-go, vs. natural = no-go, man-made = go), but we used different words for each go/no-go rule (resulting in consistent stimulus-stop mappings). Based on these findings, we proposed the automatic inhibition hypothesis: 'automatic inhibition' occurs when old no-go stimuli retrieve the stop goal when they

are repeated, and this interferes with go processing (Verbruggen & Logan, 2008b). The stimulus–stop mapping is typically consistent in the go/no-go paradigm, so automatic inhibition is likely to occur. However, automatic inhibition can also occur in the stop-signal task when the mapping is manipulated (Verbruggen & Logan, 2008b, Experiment 5).

The experiments of Verbruggen and Logan demonstrated behaviourally that response inhibition is not always an effortful or deliberate act of control. A follow-up neuroimaging study showed that the right inferior frontal gyrus, which is part of the fronto-basal-ganglia network that supports deliberate response inhibition (see above), was also activated when stimuli previously associated with stopping were presented in a stop-signal task (Lenartowicz, Verbruggen, Logan, & Poldrack, 2011). Thus, at least part of the top-down inhibition network was activated in the absence of external stop signals. However, the rIFG has been associated with a multitude of roles (e.g. attentional reorientation, context monitoring, response selection, reversal learning), thus this finding does not necessarily allow strong inferences about the underlying cognitive mechanisms.

What is learned during conditioning of inhibitory control?

What is learned during go/no-go and stop-signal tasks is still unclear. Based on Logan's Instance Theory of Automatization (1988), we hypothesized that stimuli became associated with a stop goal or stop representation in training, which impaired responding to them at test (Verbruggen et al., 2008; Verbruggen & Logan, 2008b). Like "No-US" representations (Part I), stop representations can be interpreted in different ways. First, the stop representation could be response specific. When a cue or stimulus is trained with stopping a left manual response, the stop representation would be 'stop left response' (or to be even more specific, 'stop left hand response'); but when the stimulus is trained with stopping a right response, the stop representation would be 'stop right response'. Second, the stop representation could be more general. Previously we have argued that in stop-signal tasks, a stimulus becomes associated with an abstract and general representation of going or stopping; in other words, it does not specify which specific response or motor program has to be executed or stopped (Verbruggen & Logan, 2008). The study of Giesen and Rothermund (2013) provides direct support for this general representation idea. These authors demonstrated that responding to a stimulus that was previously associated with stopping, was delayed even when the expected go response had changed. More specifically, the colour of a letter indicated whether subjects had to execute a left or right response; the identity of the letter ('D' or 'L') was irrelevant. They found that responding to a letter was slowed down if a stop signal was presented on the previous trial, regardless of the 'to-be-executed' or 'to-be-stopped' response (e.g. a green D on the prime, followed by a red D). This suggests that the stimulus-stop associations are general. Note that the 'general stop representation' idea is also indirectly supported by the observation that stopping often has general effects on the motor system (see above).

Recent work on stopping to motivationally salient stimuli suggests a third interpretation. Several studies have found that consistent pairing of food-related pictures to stopping in a go/no-go or stop-signal-paradigm reduced subsequent food consumption (e.g. Houben, 2011; Houben & Jansen, 2011; Lawrence, Verbruggen, Adams, & Chambers, 2013; Veling, Aarts, & Papies, 2011; Veling, Aarts, & Stroebe,

2012). Furthermore, a similar procedure with alcohol-related stimuli reduced alcohol-intake in the laboratory (Jones & Field, 2013) and even self-reported weekly alcohol intake of heavy drinking students (Houben, Havermans, Nederkoorn, & Jansen, 2012; but see Jones & Field, 2013). These effects could be mediated by devaluation of the stimuli that were associated with stopping (e.g. Houben et al, 2012; Kiss, Raymond, Westoby, Nobre, Eimer, 2008; Veling, Holland, and van Knippenberg, 2008). Ferrey, Frischen, and Fenske (2012) showed that stop associations not only impact on the hedonic value of the stimuli associated with stopping but also on their behavioral incentive. They paired sexually attractive images with either going or stopping in a training phase, and then asked subjects to rate the attractiveness of the images. They found that the nogo (stop) images were rated less positively than the go images. This is similar to the findings of Raymond et al., who showed that ignoring a distractor leads to its devaluation. In a second study, Ferrey et al showed that subjects were less willing to work to see the erotic images that were paired with stopping. Thus, conditioned inhibitory control may impact on the motivational value of stimuli, perhaps via creating links between the stimuli and the appetitive/aversive centres postulated by Dickinson and Dearing (1979).

Central to the 'conditioned inhibitory control' idea is the notion that the retrieval of stop representations will impair responding. However, such impairments could arise in at least two different processing stages: action selection and action execution[3]. First, in go/no-go and stop-signal tasks, subjects must select an action on each trial (Gomez, Ratcliff, Perea, 2007; Logan et al., 2014). The retrieval of stop information could interfere with selecting the appropriate 'go' action. This would be akin to 'central' interference between two competing go responses when selecting a response. This also implies that conditioned inhibitory control could be achieved via lateral local inhibitory connections between competing action options. This interference or conflict account receives some support from short-term aftereffect studies which demonstrated that stopping on the previous trial affected the stimulus-locked parietal P300, but only when the stimulus was repeated (Oldenburg et al., 2012). Response-locked motor components were not influenced, arguing against a motor locus for the effect (see also Enticott et al., 2009). Second, the retrieval of the stimulus-stop association could serve as a conditioned stop 'signal', activating the indirect or hyperdirect pathways that suppress motor output. This would be more similar to the direct, unconditional, automatic activation of an incorrect go response in interference tasks. Consistent with the motor suppression idea, Chiu, Aron, and Verbruggen (2012) showed that motor excitability was suppressed a mere 100 ms after the presentation of stimuli that were previously associated with stopping, but now required going. Of course, the two options are not exclusive. They may even rely on overlapping neural structures. The detection of conflict (defined as the competition between response options) could trigger a braking mechanism via the No-go cells of the indirect pathway or the hyperdirect pathway (see above; Frank, 2006; Ratcliff & Frank, 2012). If conflict between go and stop representations is detected early enough then this braking mechanism could account for the reduced motor excitability observed in Chiu et al. (2012). Thus, the main difference between the 'automatic suppression' account and the 'conflict' account is the trigger of the braking or stopping mechanism: the stimulus itself or the conflict caused by the retrieved information, respectively. Future

[3] In Verbruggen, Best, Bowditch, Stevens, and McLaren (2014), we discuss a third possibility, namely that attention and signal detection become conditioned.

work is required to determine how exactly stop representations influence responding in various situations.

In combination, the work above suggests that inhibitory control can be conditioned or become ‘automatized’. Dickinson and Dearing (1979) made a strong case for motivational influences and an appetitive-aversive interaction in Pavlovian conditioning. The work on conditioned inhibitory control suggests that very similar mechanisms might operate in instrumental inhibitory conditioning, despite the fact that Pavlovian and instrumental conditioning differ in many other ways (c.f. Dickinson & Balleine, 2002). In the next section, we will focus on integrating these findings and develop a theory of how ‘conditioned’ or ‘automatic’ inhibition might operate.

Integration: Inhibition and Association

Here we ask if it is possible to bring these two very different areas (animal conditioning and human cognitive psychology) together and arrive at a unified treatment of "inhibition" that would make sense in both domains. Our (somewhat tentative) answer is that it may be possible to develop an integrated approach that captures an emerging consensus in the two separate areas. This consensus revolves more around the associative structures that need to be posited to capture the notion of inhibition than the particular learning algorithms needed to operate within those structures, and so our treatment will mostly focus on the general architecture of inhibition at this point rather than exactly how it develops within this architecture (though the two issues are clearly not independent of one another).

To recap, the work reviewed in Part I (‘Conditioned Inhibition’) suggests that there is a general form of inhibition mediated via excitatory connections to appetitive/aversive centres that mutually inhibit one another, and a more specific form of inhibition that is equivalent to a direct inhibitory link to the stimulus representation (be it CS or US) in question. Both will contribute to learning, and task contexts might determine the relative contribution of the two. The work reviewed in Part II (‘Conditioned Inhibitory Control’) suggests that inhibition of responses is an integral part of executive control, but in many situations, this top-down response inhibition can become ‘automatized’. Recent work suggests that subjects learn a general form of response inhibition, which transfers between tasks. This could be mediated by the same excitatory connections to the appetitive and aversive centres that are a key component of Pavlovian conditioning. Indeed, learning to stop or not to respond to a certain stimulus not only slows responding to it (e.g. Lenartowizc et al, 2011; Neill et al, 1992; Verbruggen & Logan, 2008b) but also reduces its hedonic value and motivational incentive (e.g. Ferrey et al., 2012; Houben et al, 2012; Kiss et al, 2008; Raymond et al, 2003; Veling et al, 2008). Our interpretation of this is that when a distractor or no-go/stop stimulus becomes associated with an avoidance/aversive centre, then presentation of it will directly activate the avoidance/aversive centre, which in turn will suppress activation of the approach/appetitive centre (cf. Dickinson & Balleine, 2002). This could explain both the slower responding in a RT task and the lower hedonic values in a stimulus evaluation task using ratings.

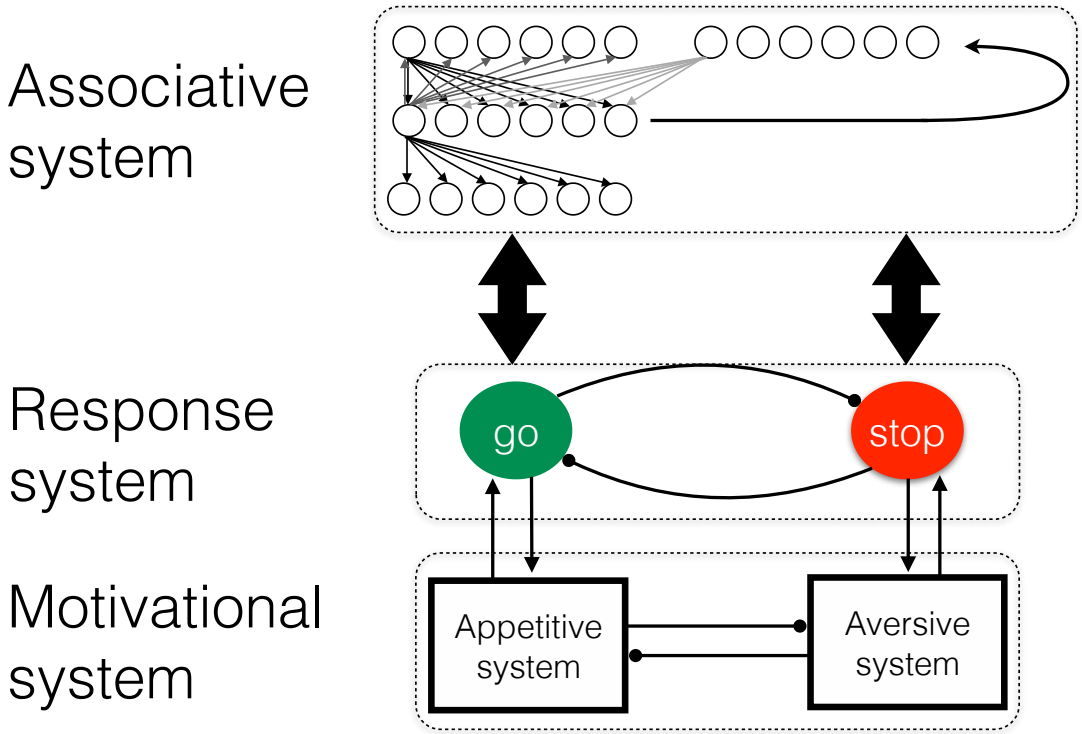

Figure 3: A model integrating associative and motivational sub-systems that would enable implementation of our proposals for conditioned inhibition. The associative system contains both an auto-associative network and recurrence giving it the ability to capture statistical regularities in the environment and between actions and outcomes. The motivational and response systems are a synthesis of Dickinson and Balleine's (2002) implementation of Konorski's proposal with an instrumental Stop/Go system along the lines proposed by Boucher et al (2007). "Direct" conditioned inhibition takes place within the associative system, and is outcome specific. "General" conditioned inhibition takes place via links from the associative system to the other systems either at the Stop/Go instrumental level or the Appetitive/Aversive Pavlovian level.

In a sense, then, we are arguing that "Go" and "Stop" are the instrumental equivalents of the Pavlovian "Good" and "Bad", and a scheme that implements this idea is shown in outline in Figure 3. Of course, Pavlovian and instrumental conditioning should not be equated entirely, as they appear to be influenced in different ways by manipulations of contexts, omission schedules (Dickinson & Balleine, 2002), and they are supported by different corticostriatal loops (for a short review, see Guitart-Masip, Duzel, Dolan, Dayan, 2014). Nevertheless, recent work suggests that Pavlovian and instrumental conditioning interact in a go/no-go task (Guitart-Masip et al, 2014). For example, in a study by Guitart-Masip et al. (2012), subjects had to learn stimulus-go/no-go contingencies. They learned them faster when correct go responses were rewarded and incorrect no-go responses were punished, than the other way around. This was attributed to a hard-wired Pavlovian equivalence between reward/punishment and approach/avoidance, respectively. The Konorskian model, as discussed in Dickinson & Balleine (2002), also links the aversive system with avoidance (withdrawal, suppression) and the appetitive system with approach (go). Therefore, it seems plausible to suggest that when subjects always have to stop their response to a specific stimulus, a link between this stimulus and the aversive/avoidance system will be created.

Despite the seemingly overwhelming evidence for a strong link between go and appetite/reward and between no-go and aversion/punishment, a few findings appear inconsistent with the no-go/aversion account. For example, some studies have shown that response inhibition might be impaired rather than enhanced when negative emotional or threatening stimuli are presented (e.g. De Houwer & Tibboel, 2010; Pessoa, Padmala, Kenzer, & Bauer, 2012; Verbruggen & De Houwer, 2007). Because these studies showed similar impairments when positive stimuli were presented, the effect of emotional and threatening stimuli has been attributed to arousal (rather than valence): arousing stimuli tend to attract attention (and are processed centrally when they high in threat), causing 'dual-task' interference. In other words, effects of arousal (attention) may have counteracted or dominated the effects of valence (positive/negative). Perhaps this is not very surprising given recent work that suggests that most of the stopping latency is occupied by afferent or sensory processes (Boucher et al., 2007; Salinas & Stanford, 2013); in other words, activation of the avoidance/aversive centre may only have a small influence on the overall SSRT, compared with the effect of arousal, because of the different time courses for the processes involved. In the study by Pessoa et al (2012, Experiment 2) in particular, the latency for activation of any aversive centre due to associations between the stimulus and some motivationally significant outcome may have been too long for it to have much effect on stopping in the stop-signal task, making any effect entirely dependent on a more cognitive appraisal of the stimulus.

So far, we have focused mostly on the link between conditioned inhibitory control and appetitive/aversive valence. But our discussion of the conditioned inhibition literature suggests that performance cannot be explained using a single inhibitory mechanism. Apart from the direct link between the CS and the appetitive/aversive centres, there is the more specific link between the CS and US (or another CS). In the case of conditioned inhibition, this link will be inhibitory. Of course, in many other situations, this link will be excitatory (as in the original work of Pavlov). It seems likely that in the context of conditioned inhibitory control, subjects can also learn associations between the representation of the go stimulus and the representation of the stop signal (Verbruggen, Best, et al., 2014). Factors such as the number and kind of stop or no-go signals could determine the relative contribution of stimulus-stimulus associations vs. stimulus-approach/avoidance associations.

Conclusion: Inhibition in Cognitive Control and Associative Learning

We have tried to provide a modern approach to the problem of inhibition that draws on many of the classic studies in the animal learning tradition that exemplify the contribution that experimental psychology can make to current issues in cognitive neuroscience. We hope that this integration of the old and the new will prove fruitful in providing a framework for future research on behavioural inhibition.

References

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences.* doi:10.1016/j.tics.2013.12.003

Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, *49A*(1), 5-28

Badry, R., Mima, T., Aso, T., Nakatsuka, M., Abe, M., Fathi, D., . . . Fukuyama, H. (2009). Suppression of human cortico-motoneuronal excitability during the stop-signal task. *Clinical Neurophysiology*, *120*, 1717-23. doi:10.1016/j.clinph.2009.06.027

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: Behavioral and neural basis of response control. *Progress in Neurobiology.* doi:10.1016/j.pneurobio.2013.06.005

Bennett, C.H., Scahill, V. L., Griffiths, D. P. and Mackintosh, N. J. (1999). The role of inhibitory associations in perceptual learning. *Learning & Behavior*, Volume 27, 333-345.

Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: Control adjustments in the stop-signal paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 392-404. doi:10.1037/a0021800

Buch, E. R., Mars, R. B., Boorman, E. D., & Rushworth, M. F. (2010). A network centered on ventral premotor cortex exerts both facilitatory and inhibitory control over primary motor cortex during action reprogramming. *The Journal of Neuroscience*, *30*, 1395-401. doi:10.1523/JNEUROSCI.4882-09.2010

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective & Behavioral Neuroscience*, *4*, 564-79.

Chambers, C. D., Garavan, H., & Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neuroscience and Biobehavioral Reviews*, *33*, 631-646. doi:10.1016/j.neubiorev.2008.08.016

Chapman, G.B. (1991) . Trial-order affect s cue interaction in contingency judgement. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 837–854.

Chiu, Y. C., & Aron, A. R. (2013). Unconsciously triggered response inhibition requires an executive setting. *Journal of Experimental Psychology: General.* doi:10.1037/a0031497

Chiu, Y. C., Aron, A. R., & Verbruggen, F. (2012). Response suppression by automatic retrieval of stimulus-stop association: Evidence from transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, *24*, 1908-18. doi:10.1162/jocn_a_00247

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the stroop effect. *Psychological Review*, *97*, 332-61.

Cotton, M.M., Goodall, G. & Mackintosh, N.J. (1982). Inhibitory conditioning resulting from a reduction in the magnitude of reinforcement. *Quarterly Journal of Experimental Psychology*, 34B, 163-180.

De Houwer, J., & Tibboel, H. (2010). Stop what you are not doing! Emotional pictures interfere with the task not to respond. *Psychonomic Bulletin & Review*, *17*(5), 699-703. doi:10.3758/PBR.17.5.69

Dickinson, A. (1980). *Contemporary animal learning theory*. CUP.

Dickinson, A., & Balleine, B. (2002). The role of learning in the operation of motivational systems. In R. Gallistel (Ed.), *Stevens' handbook of experimental psychology* (3 ed., Vol. 3, pp. 497- 534). New York: John Wiley & Sons.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, 49B, 60–80.

Dickinson, A. & Dearing, M.F. (1979). Appetitive-aversive interactions and inhibitory processes. In *Mechanisms of Learning and Motivation*, A. Dickinson and R.A. Boakes (Eds.) LEA, 203-32.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475-83.

Enticott, P. G., Bradshaw, J. L., Bellgrove, M. A., Upton, D. J., & Ogloff, J. R. (2009). Stop task after-effects: The extent of slowing during the preparation and execution of movement. *Experimental Psychology*, *56*, 247-51. doi:10.1027/1618-3169.56.4.247

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

Frank, M. J. (2006). Hold your horses: A dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw*, *19*, 1120-36. doi:10.1016/j.neunet.2006.03.006

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology. General*, *133*, 101-35. doi:10.1037/0096-3445.133.1.101

Fritz, J., & Dreisbach, G. (2013). Conflicts as aversive signals: Conflict priming increases negative judgments for neutral stimuli. *Cognitive, Affective & Behavioral Neuroscience*, *13*(2), 311-7. doi:10.3758/s13415-012-0147-

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology. General*, *136*(3), 389-413. doi:10.1037/0096-3445.136.3.38.

Graham S (1999). Retrospective revaluation and inhibitory associations: does perceptual learning modulate our perception of the contingencies between events? *Quarterly Journal of Experimental Psychology*, 52B, 159-185.

Graham S., Jie H.L., Minn C.H., McLaren I.P.L., Wills A.J., (2011). Simultaneous Backward Conditioned Inhibition and Mediated Conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, Vol 37(2), Apr 2011, 241-245.

Greenhouse, I., Oldenkamp, C. L., & Aron, A. R. (2011). Stopping a response has global or non-global effects on the motor system depending on preparation. *Journal of Neurophysiology.* doi:10.1152/jn.00704.2011

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences.* doi:10.1016/j.tics.2014.01.00

Hall, G. and Pearce, J.M. (1979). Latent inhibition of a CS during CS-US parings. *Journal of Experimental Psychology: Animal Behaviour Processes*, 5, 31-43.

Harris, J. A., Kwok, D. W. S., & Andrew, B. J. (in press). Conditioned inhibition and reinforcement rate. *Journal of Experimental Psychology: Animal Learning & Cognition*.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 175-204). San Diego, CA: Academic Press.

Herd, S. A., Banich, M. T., & O'Reilly, R. C. (2006). Neural mechanisms of cognitive control: An integrative model of stroop task performance and FMRI data. *Journal of Cognitive Neuroscience*, *18*, 22-32. doi:10.1162/089892906775250012

Houben, K. (2011). Overcoming the urge to splurge: Influencing eating behavior by manipulating inhibitory control. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*, 384-388. doi:10.1016/j.jbtep.2011.02.008

Houben, K., & Jansen, A. (2011). Training inhibitory control. A recipe for resisting sweet temptations. *Appetite*, *56*, 345-9. doi:10.1016/j.appet.2010.12.017

Houben, K., Havermans, R. C., Nederkoorn, C., & Jansen, A. (2012). Beer à no-go: Learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction*, *107*, 1280-7. doi:10.1111/j.1360-0443.2012.03827.x

Jones, A., & Field, M. (2013). The effects of cue-specific inhibition training on alcohol consumption in heavy social drinkers. *Experimental and Clinical Psychopharmacology*, *21*, 8-16. doi:10.1037/a0030683

Jones, F., Wills, A.J., and McLaren, I.P.L. **(**1998). Perceptual categorisation: connectionist modelling and decision rules*. Quarterly Journal of Experimental Psychology*, 51B, 33-58.

Kiss, M., Raymond, J. E., Westoby, N., Nobre, A. C., & Eimer, M. (2008). Response inhibition is linked to emotional devaluation: Behavioural and electrophysiological evidence. *Frontiers in Human Neuroscience*, *2*, 13. doi:10.3389/neuro.09.013.200

Klein, P. A., Petitjean, C., Olivier, E., & Duque, J. (2014). Top-down suppression of incompatible motor activations during response selection under conflict. *NeuroImage*, *86*(1), 138–149. doi:10.1016/j.neuroimage.2013.08.00

Konorski, J. (1948). *Conditioned Reflexes and Neuron Organisation.* Cambridge: Cambridge University Press.

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

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility - a model and taxonomy. *Psychological Review*, *97*, 253-270.

Kremer, E. F. (1978). The Rescorla–Wagner model: Losses of associative strength in compound conditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 4, 22–36.

Lawrence, N. S., Verbruggen, F., Adams, R. C., & Chambers, C. D. (2013). Stopping to food cues reduces snack food intake: Effects of cue specificity, control conditions and individual differences. *In Preparation.*

Lenartowicz, A., Verbruggen, F., Logan, G. D., & Poldrack, R. A. (2011). Inhibition-related activation in the right inferior frontal gyrus in the absence of inhibitory cues. *Journal of Cognitive Neuroscience*, *23*, 3388-99. doi:10.1162/jocn_a_00031

Le Pelley, M.E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *Quarterly Journal of Experimental Psychology*, 57B, 193-243.

Le Pelley, M.E., Cutler, D.L., and McLaren, I.P.L. (2000). Retrospective Effects in Human Causality Judgment. *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*, NJ, LEA.

Le Pelley, M. E., & McLaren, I. P. L. (2001). Retrospective revaluation in humans: Learning or memory? *Quarterly Journal of Experimental Psychology,* 54B, 311-352.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica*, *60*, 193-210

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492-527

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, *91*, 295-327.

Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E. J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, *121*(1), 66-95. doi:10.1037/a0035230

Lotz, A., Uengoer, M., Koenig, S., Pearce, J. M. and Lachnit, H. (2012). An exploration of the feature-positive effect in adult humans. *Learning & Behavior*, 40(2), 222-230.

McLaren, I.P.L. and Dickinson, A. (1990). The conditioning connection. *Phil. Trans. R. Soc. Lond.* B, 329, 179-186

McLaren, I. P. L., Forrest, C.L., and McLaren, R.P. (2012). Elemental representation and configural mappings: Combining elemental and configural theories of associative learning. *Learning and Behavior*, 40 (3), 320-33.

McLaren, I. P. L. & Mackintosh, N. J. (2002).Associative learning and elemental representation: II. Generalization and discrimination. Animal Learning & Behavior, 30, 3, 177-200.

McLaren, I.P.L. and Mackintosh, N.J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. Animal Learning

and Behavior, *38(3),* 211-246.

McLaren, I.P.L., Kaye, H. and Mackintosh, N.J. (1989). An associative theory of the representation of stimuli: applications to perceptual learning and latent inhibition. In R.G.M. Morris (Ed.) Parallel Distributed Processing - Implications for Psychology and Neurobiology. Oxford. OUP.

Mackintosh, N.J. 1974. *The Psychology of Animal Learning*. London. Academic Press.

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

MacLeod, C. M., Dodd, M. D., Sheard, E. D., Wilson, D. E., & Bibi, U. (2003). In opposition to inhibition. In *Psychology of learning and motivation: Advances in research and theory, vol 43* (Vol. 43, pp. 163-214).

Majid, D. S., Cai, W., George, J. S., Verbruggen, F., & Aron, A. R. (2012). Transcranial magnetic stimulation reveals dissociable mechanisms for global versus selective corticomotor suppression underlying the stopping of action. *Cerebral Cortex*, *22*, 363-71. doi:10.1093/cercor/bhr112

Mayr, S., & Buchner, A. (2007). Negative priming as a memory phenomenon. *Zeitschrift Für Psychologie / Journal of Psychology*, *215*, 35-51. doi:10.1027/0044-3409.215.1.35

Miller, R. R., & Schachtman, T.R. (1985). Conditioning context as an associative baseline: Implications for response generation and the nature of conditioned inhibition. In R.R. Miller & N.E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 51-88). Hillsdale, NJ: Erlbaum.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. doi:10.1006/cogp.1999.0734

Mostofsky, S. H., & Simmonds, D. J. (2008). Response inhibition and response selection: Two sides of the same coin. *Journal of Cognitive Neuroscience*, *20*, 751-761.

Mundy, M.E., Dwyer, D.M., & Honey, R.C. (2006). Inhibitory associations contribute to perceptual learning in humans. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *32*, 178-184.

Neill, W. T., & Valdes, L. A. (1992). Persistence of negative priming - steady-state or decay. *Journal of Experimental Psychology: Learning Memory and Cognition*, *18*, 565-576. Retrieved from Google Scholar.

Neill, W. T., Valdes, L. A., Terry, K. M., & Gorfein, D. S. (1992). Persistence of negative priming: II. Evidence for episodic trace retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition*, *18*, 993-1000.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*, 220-246. doi:10.1037//0033-2909.126.2.220

Oldenburg, J. F., Roger, C., Assecondi, S., Verbruggen, F., & Fias, W. (2012).

Repetition priming in the stop signal task: The electrophysiology of sequential effects of stopping. *Neuropsychologia.* doi:10.1016/j.neuropsychologia.2012.08.014

Pavlov, I.P. (1927). *Conditioned Reflexes*. OUP.

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

Pessoa, L., Padmala, S., Kenzer, A., & Bauer, A. (2012). Interactions between cognition and emotion during response inhibition. *Emotion*, *12*(1), 192-7. doi:10.1037/a002410

Raymond, J.E., Fenske, M.J., & Tavassoli, N.T. (2003). Selective attention determines emotional responses to novel visual stimuli. *Psychological Science*, 14, 537–542.

Rescorla, R.A. (1969). Conditioned inhibition of fear resulting from negative CS-US contingencies. *Journal of Comparative and Physiological Psychology*, 67, 504-9.

Rescorla, R.A. (1970). Reduction in the Effectiveness of Reinforcement after Prior Excitatory Conditioning. Learning and Motivation, 1, 372-81.

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

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. *Attention and Performance XIX: Common Mechanisms in Perception and Action*, 494-519.

Rieger, M., & Gauggel, S. (1999). Inhibitory after-effects in the stop signal paradigm. *British Journal of Psychology*, *90*, 509-518.

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 482-95. doi:10.1037/0278-7393.31.3.482

Schall, J. D., & Godlove, D. C. (2012). Current advances and pressing problems in studies of stopping. *Current Opinion in Neurobiology.* doi:10.1016/j.conb.2012.06.002

Smittenaar, P., Guitart-Masip, M., Lutti, A., & Dolan, R. J. (2013). Preparing for selective inhibition within frontostriatal loops. *The Journal of Neuroscience*, *33*, 18087-97. doi:10.1523/JNEUROSCI.2167-13.2013

Swick, D., Ashley, V., & Turken, A. U. (2011). Are the neural correlates of stopping and not going identical? Quantitative meta-analysis of two response inhibition tasks. *NeuroImage*, *56*, 1655-1665. doi:10.1016/j.neuroimage.2011.02.070

van Gaal, S., Lamme, V. A., Fahrenfort, J. J., & Ridderinkhof, K. R. (2011). Dissociable brain mechanisms underlying the conscious and unconscious control of behavior. *Journal of Cognitive Neuroscience*, *23*, 91-105. doi:10.1162/jocn.2010.21431

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. (2010).

Unconscious activation of the prefrontal no-go network. *The Journal of Neuroscience*, *30*, 4143-50. doi:10.1523/JNEUROSCI.2992-09.2010

van Gaal, S., Ridderinkhof, K. R., van den Wildenberg, W. P., & Lamme, V. A. (2009). Dissociating consciousness from inhibitory control: Evidence for unconsciously triggered response inhibition in the stop-signal task. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1129-39. doi:10.1037/a0013551

Van Hamme, L.J., & Wasserman, E.A. (1994). Cue competition in causality judgements: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, 25, 127–151.

van den Wildenberg, W.P.M., Wylie, S.A., Forstmann, B.U., Burle, B., Hasbroucq, T. & Ridderinkhof, K.R. (2010). To head or to heed ? Beyond the surface of selective action inhibition : A review. *Frontiers in Human Neuroscience,4*: 222, doi: 10.3389/fnhum.2010.00222.

van Steenbergen, H., Band, G. P., & Hommel, B. (2009). Reward counteracts conflict adaptation. Evidence for a role of affect in executive control. *Psychological Science*, *20*(12), 1473-7. doi:10.1111/j.1467-9280.2009.02470.

Veling, H., Holland, R. W., & van Knippenberg, A. (2008). When approach motivation and behavioral inhibition collide: Behavior regulation through stimulus devaluation. *Journal of Experimental Social Psychology*, *44*(4), 1013-1019. doi:10.1016/j.jesp.2008.03.00

Veling, H., Aarts, H., & Papies, E. K. (2011). Using stop signals to inhibit chronic dieters' responses toward palatable foods. *Behaviour Research and Therapy*, *49*, 771-80. doi:10.1016/j.brat.2011.08.005

Veling, H., Aarts, H., & Stroebe, W. (2012). Using stop signals to reduce impulsive choices for palatable unhealthy foods. *British Journal of Health Psychology*. doi:10.1111/j.2044-8287.2012.02092.x

Verbruggen, F., Aron, A. R., Stevens, M. A., & Chambers, C. D. (2010). Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 13966-71. doi:10.1073/pnas.1001957107

Verbruggen, F., Best, M., Bowditch, W., Stevens, T., McLaren, I.P.L. (2014). The inhibitory control reflex. *Manuscript under revision*

Verbruggen, F., & De Houwer, J. (2007). Do emotional stimuli interfere with response inhibition? Evidence from the stop signal paradigm. *Cognition & Emotion*, *21*(2), 391-403. doi:10.1080/02699930600625508

Verbruggen, F., & Logan, G. D. (2008a). Long-term aftereffects of response inhibition: Memory retrieval, task goals, and cognitive control. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1229-35. doi:10.1037/0096-1523.34.5.1229

Verbruggen, F., & Logan, G. D. (2008b). Automatic and controlled response inhibition: Associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology. General*, *137*, 649-72. doi:10.1037/a0013170

Verbruggen, F., & Logan, G. D. (2008c). Response inhibition in the stop-signal

paradigm. *Trends in Cognitive Sciences*, *12*, 418-24. doi:10.1016/j.tics.2008.07.005

Verbruggen, F., & Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience and Biobehavioral Reviews*, *33*, 647-61. doi:10.1016/j.neubiorev.2008.08.014

Verbruggen, F., Logan, G. D., Liefooghe, B., & Vandierendonck, A. (2008). Short-term aftereffects of response inhibition: Repetition priming or between-trial control adjustments? *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 413-26. doi:10.1037/0096-1523.34.2.413

Verbruggen, F., McLaren, I. P. L., & Chambers, C. D. (2014). Banishing the control homunculi in studies of action control and behaviour change. *Perspectives on Psychological Science*, Advance online publication. Retrieved from http://hdl.handle.net/10871/14582

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, *46*, 361-413. doi:10.1016/S0010-0285

Waszak, F., Hommel, B., & Allport, A. (2004). Semantic generalization of stimulus-task bindings. *Psychonomic Bulletin & Review*, *11*, 1027-33.

Waszak, F., Hommel, B., & Allport, A. (2005). Interaction of task readiness and automatic retrieval in task switching: Negative priming and competitor priming. *Memory & Cognition*, *33*, 595-610.

Wiecki, T. V., & Frank, M. J. (2013). A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychological Review*, *120*, 329-55. doi:10.1037/a0031542

Zimmer-Hart C.L. and Rescorla, R.A. (1974). Extinction of Pavlovian conditioned inhibition. *Journal of Comparative and Physiological Psychology*, 86, 837-45.