



SAPIENZA
UNIVERSITÀ DI ROMA

FACOLTÀ DI MEDICINA E PSICOLOGIA

CORSO DI NEUROSCIENZE COGNITIVE E RIABILITAZIONE PSICOLOGICA

Tesi di Laurea Magistrale

**Pavlovian-instrumental transfer: un
modello computazionale bayesiano con
cause latenti**

Laureando

Emilio Cartoni

Relatore

Prof. Stefano Puglisi-Allegra

Relatore Esterno

Prof. Gianluca Baldassarre

Correlatore

Prof. Fabio Presaghi

Anno Accademico 2012–2013

Indice

Ringraziamenti	iii
1 Introduzione	1
2 Pavlovian-instrumental transfer	3
2.1 Che cos'è il PIT	3
2.2 PIT specifico e PIT generale: il paradigma	4
2.3 PIT specifico e PIT generale: basi neurali	5
2.4 Interazioni tra PIT ed altre variabili	6
3 Ipotesi: tre principi del comportamento goal-directed	10
3.1 Spiegare il PIT	10
3.2 Tre principi: contesto, efficacia, utilità	16
4 Reti bayesiane e modelli con variabili latenti	19
4.1 Modelli probabilistici	19
4.2 Reti bayesiane	20
4.3 Apprendimento pavloviano e cause latenti: il modello di Courville	22
5 Modello del PIT	25
5.1 Struttura del modello	25
5.2 Apprendimento del modello	27
5.3 Effetto PIT	31
6 Conclusioni	34

Elenco delle figure

2.1	Paradigma PIT completo	4
3.1	Modello associativo-cibernetico	11
3.2	Dati Corbit et al. 2005 - lesioni amigdala	12
3.3	Dati Corbit et al. 2007	14
3.4	Dati Hall et al. 2001 - lesioni <i>nucleus accumbens</i>	15
4.1	Esempio di rete bayesiana	20
4.2	Modello <i>goal-directed</i> di Solway & Botvinick (2012)	21
4.3	Modello pavloviano con cause latenti	23
4.4	Modello pavloviano di Courville et al. (2005)	23
5.1	Modello PIT: i nodi	26
5.2	Modello PIT: distribuzioni a posteriori	29
5.3	Modello PIT: fase Pavloviana	30
5.4	Modello PIT: fase strumentale	31
5.5	Modello PIT: fase di test	32

Ringraziamenti

Ringrazio Gianluca Baldassarre per avermi dato l'opportunità di intraprendere questo lavoro all'interno del gruppo LOCEN dell'ISTC-CNR.

1 Introduzione

Questa tesi si propone come obiettivo di formulare un modello del *Pavlovian-instrumental transfer*, ossia dell'interazione tra condizionamento pavloviano e condizionamento strumentale. In breve, il *Pavlovian-instrumental transfer* (PIT) è definibile come la capacità di uno stimolo associato pavlovianamente ad un certo rinforzo di potenziare la risposta strumentale verso lo stesso od un altro rinforzo. Ad esempio, udire un suono associato alla comparsa di cibo può aumentare la frequenza di pressione di una leva diretta ad ottenere lo stesso od un altro cibo.

Negli ultimi anni, sono stati fatti importanti progressi nell'individuare le aree del cervello implicate nel PIT. In particolare, lo stesso effetto PIT è stato suddiviso in due tipi: PIT specifico e PIT generale, ognuno dei quali ha come base un diverso substrato neurale. Entrambi gli effetti, PIT specifico e PIT generale, sono mediati sia dall'amigdala che dal *nucleus accumbens*, ma mentre il PIT specifico coinvolge l'amigdala basolaterale¹ e la parte *shell* del *nucleus accumbens*, il PIT generale coinvolge l'amigdala centrale e la parte *core* del *nucleus accumbens* (Corbit & Balleine, 2005, 2011). Nonostante questi due effetti siano stati localizzati in termini di aree², non sono però ancora noti i meccanismi neurali tramite cui queste strutture danno luogo all'interazione tra condizionamento pavloviano e condizionamento strumentale.

Al di là del livello del meccanismo neurale, anche a livello funzionale la situazione non è chiara. Le spiegazioni in termini di catene *stimulus-outcome-response* ($S \rightarrow O \rightarrow R$), in cui lo stimolo condizionato elicitava la rappresentazione di un outcome (il cibo) che a sua volta richiama una risposta strumentale, non sono pienamente soddisfacenti: non spiegano infatti come mai PIT specifico e PIT generale tendano ad escludersi a vicenda (nonostante siano apparentemente meccanismi paralleli) o perché il PIT generale non si verifichi nei confronti di alcune risposte strumentali.

¹Usiamo qui ed in seguito le espressioni 'amigdala basolaterale' ed 'amigdala centrale' per indicare rispettivamente i nuclei del complesso basolaterale dell'amigdala ed il nucleo centrale dell'amigdala.

²Vi sono indizi anche del coinvolgimento di altre aree, quali ad esempio la corteccia orbitofrontale (Ostlund & Balleine, 2007), lo striato dorsale (Corbit & Janak, 2010) o l'area tegmentale ventrale (El-Amamy & Holland, 2007) ma la maggior parte dei dati al momento sono su amigdala e *nucleus accumbens*.

1 Introduzione

L'obiettivo di questa tesi è di fornire una spiegazione a livello funzionale di tutti gli aspetti del *Pavlovian-instrumental transfer* all'interno di un quadro unitario di tipo probabilistico bayesiano. In particolare, utilizzeremo un modello generativo bayesiano con variabili causali latenti: sosteniamo che i soggetti all'interno di un paradigma PIT (solitamente dei ratti³), non solo apprendono le relazioni tra le variabili osservabili (es: leve, suoni, cibo) ma assumono l'esistenza di cause nascoste sottostanti alla co-occorrenza di alcune variabili. Le inferenze sulla presenza o meno di tali cause nascoste producono i vari effetti PIT.

Tramite questo modello, proveremo a fornire una spiegazione funzionale degli aspetti del PIT che sia più completa e più specifica delle spiegazioni esistenti. Dal modello deriveremo inoltre nuove predizioni da testare con nuovi esperimenti. Poiché l'attività dei neuroni può essere pensata in termini di computazione probabilistica, l'uso di un modello probabilistico bayesiano può offrire inoltre la possibilità di un'immediata conversione dal livello funzionale al livello dei meccanismi neurali (es. Solway & Botvinick, 2012).

L'importanza dell'avere un modello con cui comprendere il PIT sta nel fatto che questo fenomeno si situa 'a cavallo tra il mondo pavloviano e quello strumentale'. Il PIT è quindi una finestra di osservazione tramite cui ottenere indizi sia sul funzionamento del condizionamento pavloviano sia sul funzionamento del condizionamento strumentale. Le strutture che sottostanno al PIT, ossia l'amigdala ed il *nucleus accumbens* sono strutture chiave coinvolte nell'assegnazione di valori e nel comportamento motivato in generale (Balleine & Killcross, 2006; Humphries & Prescott, 2010). Oltre alla rilevanza di questo argomento come ricerca di base sul funzionamento normale, vi sono anche importanti risvolti sul fronte della comprensione di aspetti patologici. Il PIT è infatti chiamato in causa anche nella spiegazione di alcuni aspetti della tossicodipendenza, ad esempio nello spiegare come gli stimoli associati al consumo di droghe possano elicitarne la ricerca della sostanza (Robinson & Berridge, 2003). È stato infatti provato in modelli animali che l'uso di sostanze quali anfetamine, cocaina o alcool può potenziare o alterare il fenomeno PIT (Wyvell & Berridge, 2001; LeBlanc et al., 2012; Corbit & Janak, 2007a). Avere un modello del funzionamento del PIT potrà quindi risultare utile anche nello studio della tossicodipendenza.

³Esistono comunque studi anche con soggetti umani, peraltro con risultati simili anche dal punto di vista della localizzazione neurale (Prévost et al., 2012).

2 Pavlovian-instrumental transfer

2.1 Che cos'è il PIT

Il *Pavlovian-instrumental transfer* (PIT) è la capacità di uno stimolo condizionato pavlovianamente (CS) di influenzare il comportamento strumentale. In particolare, uno stimolo condizionato associato ad una certa ricompensa (es: cibo), può incrementare la frequenza di un comportamento strumentale (es: premere una leva) diretto all'ottenimento della stessa ricompensa o di una diversa (es: stesso cibo o un cibo diverso). A seconda che la ricompensa associata al CS sia o meno la stessa associata al comportamento strumentale avremo due diversi tipi di PIT: specifico (stessa ricompensa) o generale (ricompensa diversa).

Il paradigma con cui viene misurato il PIT si compone di tre fasi: una fase di condizionamento pavloviano, una fase di apprendimento strumentale ed infine la fase di test del PIT. Nella fase di condizionamento pavloviano, un certo stimolo (es: un suono) viene associato all'arrivo della ricompensa (es: cibo). Nella fase strumentale viene appresa una risposta, quale il premere una leva, che permette di ottenere a sua volta il cibo. Infine, nella terza fase, per la prima volta sono contemporaneamente presenti il suono e la leva e si misura quindi se la presenza del suono aumenta la frequenza di risposta rispetto a quando il suono è assente. Solitamente, quello che accadrà è che la leva sarà premiata più frequentemente durante il suono, nonostante non vi sia stato precedentemente alcun apprendimento della pressione della leva in presenza del suono. Durante la fase di test sia il suono che la leva sono presentati in estinzione, ossia senza il cibo, per evitare che l'effetto possa essere dovuto ad un apprendimento per rinforzo durante quest'ultima fase.

Le prime due fasi, pavloviana e strumentale, non sono necessariamente in questo ordine: è possibile effettuare il condizionamento pavloviano sia prima (Corbit & Balleine, 2011) sia dopo l'apprendimento strumentale (Wyvell & Berridge, 2001); in alcuni casi i due apprendimenti sono anche stati fatti "in parallelo" alternando sessioni pavloviane a sessioni strumentali (Blundell et al., 2001). Allo stesso modo, vi sono più varianti anche

per quanto riguarda l'utilizzo di uno o più CS con differenti ricompense e di una o più risposte strumentali.

2.2 PIT specifico e PIT generale: il paradigma

La versione “più completa” del paradigma PIT, ossia capace di distinguere tra PIT specifico e PIT generale è stata introdotta da Corbit & Balleine (2005) (vedi figura 2.1). Dato che la maggior parte degli studi sul PIT sono stati effettuati sui ratti, descriveremo il paradigma da questo punto di vista.

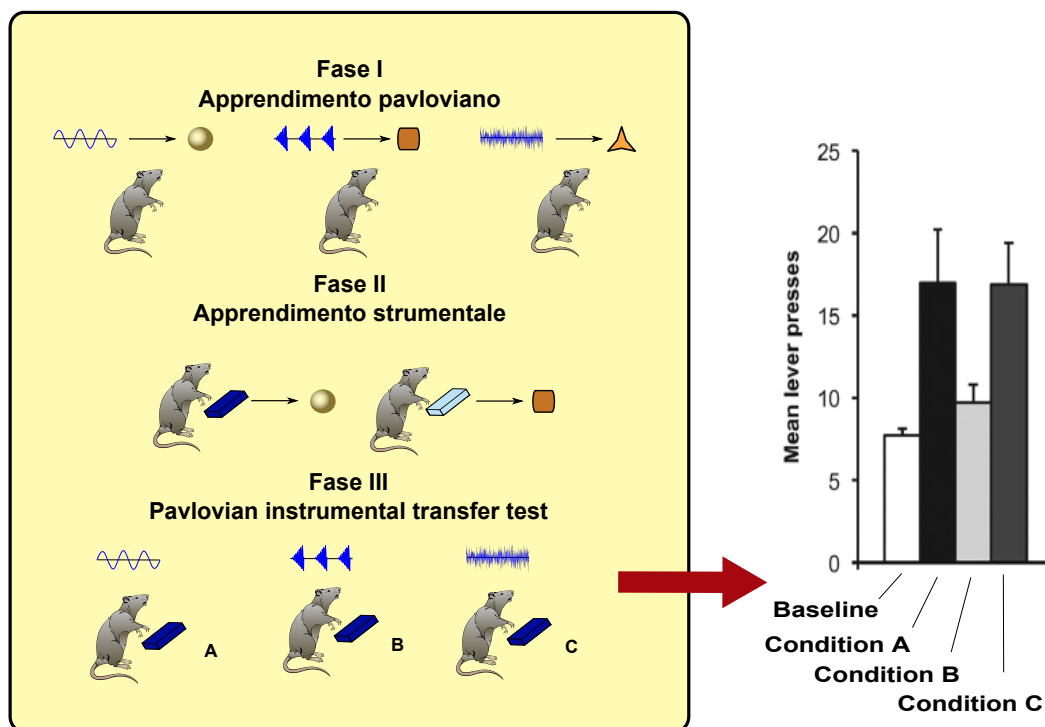


Figura 2.1: Paradigma PIT che distingue il PIT specifico dal PIT generale. L'istogramma, che mostra i risultati tipici, è tratto da Corbit & Balleine (2011).

Nella prima fase pavloviana ai ratti vengono fatti associare tre suoni diversi all'arrivo di tre cibi diversi. Successivamente, nella fase strumentale, i ratti vengono addestrati, in sessioni separate, a premere due leve. In ogni singola sessione è presente una sola leva, alla cui pressione corrisponde l'arrivo di un certo cibo, diverso a seconda di quale delle due leve sia presente. Questi due cibi della fase strumentale corrispondono a due dei cibi già usati nella fase pavloviana. Il programma di rinforzo delle leve non è continuo ed è solitamente di tipo *random ratio*, ad es: RR20, cioè in media una pressione su

venti viene ricompensata. Nella terza fase, viene infine testato l'effetto PIT. Al ratto viene nuovamente data la possibilità di premere una delle leve, in estinzione, ma questa volta vengono fatti ascoltare, ad intervalli ed ognuno separatamente, i tre diversi suoni, anch'essi in estinzione. Si vengono quindi a creare quattro condizioni a seconda se, oltre alla leva, sia presente uno dei tre suoni (condizioni A, B, C in figura 2.1) o nessun suono (*baseline*). Nella **condizione *baseline***, il ratto si trova di fronte alla leva senza nessun suono: tenderà quindi a premerla con una certa frequenza, a causa dell'apprendimento strumentale precedente. Nella **condizione A (PIT specifico)**, il ratto si trova di fronte alla leva e contemporaneamente sente il suono associato allo stesso cibo della leva. In questa condizione, il ratto premerà più frequentemente la leva (PIT *specifico*). Nella **condizione C (PIT generale)**, il ratto sentirà il suono associato ad un cibo che non è stato usato nella fase strumentale. Anche in questo caso, il ratto premerà più frequentemente la leva, ma il PIT sarà di tipo *generale*. Infine, nella **condizione B**, il ratto sentirà il suono associato al cibo che è stato usato come ricompensa per l'altra leva, non presente in quel momento. In quest'ultima condizione non vi è, in genere¹, alcun effetto PIT e la frequenza di risposta rimane quindi a livello della *baseline*.

2.3 PIT specifico e PIT generale: basi neurali

Sia nella condizione 'PIT specifico' (in cui CS e pressione della leva sono associati allo stesso cibo), sia nella condizione 'PIT generale' (in cui CS e pressione della leva sono associati a cibi diversi), si verifica un aumento della frequenza del comportamento strumentale di pressione della leva. Questo aumento di risposta è spesso di entità paragonabile (vedi istogramma in figura 2.1) ma è generato da processi neurali diversi. È stato infatti dimostrato che lesioni a carico dell'amigdala basolaterale o del *nucleus accumbens shell* eliminano l'effetto PIT solo nella condizione di 'PIT specifico', mentre lesioni all'amigdala centrale o del *nucleus accumbens core*, eliminano selettivamente l'effetto PIT solo nella condizione di 'PIT generale'. Questa dissociazione è stata confermata anche tramite inattivazione farmacologica del *core* e dello *shell* (Corbit & Balleine, 2011) ed anche tramite una procedura di disconnessione tra amigdala basolaterale e *nucleus accumbens core* e *shell* (Shiflett & Balleine, 2010). Vi sono anche alcuni dati relativi ad altre aree, in particolare relativi alla corteccia orbitofrontale (Ostlund & Balleine, 2007), allo striato dorsale (Corbit & Janak, 2007b, 2010) ed a due aree dopaminergiche: area tegmentale ventrale (Corbit et al., 2007; El-Amamy & Holland, 2007) e *substan-*

¹In altri casi si ha un effetto minore sia rispetto alla condizione di PIT specifico che di PIT generale.

tia nigra (El-Amamy & Holland, 2007). Infine, sono stati riportati dati in merito alla relazione tra PIT e trasmissione dopaminergica. In Lex & Hauber (2008) è stato dimostrato che antagonisti dei recettori D1 e D2, iniettati nel *nucleus accumbens*, sia *core* che *shell*, possono ridurre o eliminare il PIT. La relazione tra PIT e dopamina non sembra però estendersi allo striato dorsomediale, in quanto Pielock et al. (2011) non riportano praticamente alcun effetto della deplezione di dopamina né nello striato dorsomediale anteriore né dorsomediale posteriore.

2.4 Interazioni tra PIT ed altre variabili

Il PIT può essere influenzato da numerose variabili sperimentali, tra cui: il tipo di programma di rinforzo strumentale, la durata dell'apprendimento pavloviano, l'uso di stimoli diffusi o localizzati, la svalutazione del rinforzo e l'assunzione di droghe.

2.4.1 Programmi di rinforzo strumentali

I programmi di rinforzo usati nell'ambito del PIT sono generalmente di due tipi: *random-ratio* (RR) e *variable interval* (VI). Nel primo caso, ogni pressione della leva ha una certa probabilità di ottenere il rinforzo. Il numero medio di pressioni necessarie per ottenere un singolo rinforzo dà il nome al programma usato: ad esempio, RR5 indica che in media verrà rinforzata una pressione su 5, ossia ogni pressione ha il 20% di possibilità di ottenere un rinforzo. Nel caso dei programmi ad intervallo variabile (VI) invece, viene rinforzata la prima pressione eseguita dopo un certo intervallo variabile dall'ultimo rinforzo. Ad esempio, in un programma VI45, verrà rinforzata la prima pressione dopo in media 45 secondi dall'ultimo rinforzo. I programmi ad intervallo tendono a favorire lo sviluppo di *habits*, mentre nei programmi *random-ratio* il comportamento rimane più a lungo *goal-directed*, ossia sensibile al valore della ricompensa e non abituale (Yin & Knowlton, 2006). È noto inoltre che l'uso di due leve con due ricompense diverse può fare in modo che il comportamento si mantenga *goal-directed* anche con programmi ad intervallo. In genere la maggior parte degli esperimenti sul PIT sono stati eseguiti con programmi ad intervallo, con l'importante eccezione però del gruppo di Balleine e colleghi, che usano programmi RR.

All'inizio di questo secolo, era sorta una contraddizione nella letteratura: Blundell et al. (2001) riportava che la lesione dell'amigdala basolaterale aveva impatto sul PIT mentre Hall et al. (2001) e Holland & Gallagher (2003) riportavano che lesioni della

amigdala centrale e non dell'amigdala basolaterale impedivano il verificarsi del PIT. Corbit & Balleine (2005) hanno quindi riconciliato la questione mostrando, come già detto, che esistono due tipi di PIT, uno che dipende dall'amigdala basolaterale (specifico) ed uno che dipende dall'amigdala centrale (generale). Resta da capire quale sia il motivo per cui in Hall et al. (2001) ed in Holland & Gallagher (2003) il PIT evocato dalla procedura sperimentale era di tipo generale. Hall et al. (2001) e Holland & Gallagher (2003) avevano infatti usato una sola leva ed una sola ricompensa identica per la fase pavloviana e strumentale. Ci si aspetterebbe quindi, visto che la ricompensa è la stessa, un PIT specifico e non generale. Eppure il fatto che l'effetto PIT trovato fosse eliminato da lesioni all'amigdala centrale indica chiaramente che il PIT era di tipo generale. Solitamente, viene semplicemente detto che l'uso di una sola leva ed una sola ricompensa (rispetto all'uso di più leve e ricompense) 'favorisce' l'espressione del PIT generale (Corbit & Balleine, 2011). La nostra idea è invece che non sia tanto una questione di numero di leve, ma che la spiegazione risieda nell'utilizzo di una procedura che evochi o meno un comportamento di tipo *habitual* anziché *goal-directed*. L'utilizzo di programmi ad intervallo con una sola leva in Hall et al. (2001) e in Holland & Gallagher (2003) avrebbe quindi portato ad un comportamento *habitual*, mentre in Blundell et al. (2001) l'uso di due leve avrebbe mantenuto il comportamento *goal-directed* nonostante abbia anch'egli usato un programma ad intervallo. Nei comportamenti *habitual*, elicitati da programmi ad intervallo con una sola leva, la risposta non è più legata al valore della ricompensa. Se svalutiamo la ricompensa, ad esempio saziando il ratto, esso continuerà comunque a premere la leva, 'per abitudine'. Nei comportamenti *habitual* non vi è quindi una valutazione delle conseguenze dell'azione, ma l'azione è invece semplicemente elicitata dagli stimoli presenti: vedo la leva, la premo (stimolo-risposta). Poiché nella nostra ipotesi, come discuteremo più avanti, il PIT specifico è legato alla valutazione delle conseguenze dell'azione, è lecito aspettarsi che non si verifichi per comportamenti *habitual*, in cui questa valutazione non viene fatta. Il PIT generale invece, che ha più un carattere motivazionale generale, non legato necessariamente alle conseguenze dell'azione, potrebbe invece ancora essere evocato anche una volta che la risposta sia diventata abituale, al posto del PIT specifico.

2.4.2 Apprendimento pavloviano: durata e stimoli

Un altro fattore che può influenzare il PIT è la durata dell'apprendimento pavloviano. All'aumentare della durata dell'apprendimento pavloviano, diminuisce l'effetto PIT rilevabile in fase di test a causa di un effetto di competizione delle risposte (Holmes et al.,

2010). All'aumentare del apprendimento pavloviano, il ratto è infatti sempre più 'abituato' a rispondere al suono con un comportamento di avvicinamento al deposito del cibo. Questo comportamento di avvicinamento può competere con il comportamento di pressione della leva, in quanto ovviamente il ratto può eseguire solo una delle due azioni alla volta. Più è forte il comportamento 'pavloviano' di avvicinamento al deposito e meno sarà rilevabile l'effetto PIT, in quanto in presenza del suono il ratto tenderà ad andare al deposito anziché premere la leva più frequentemente.

La competizione tra risposte pavloviane e strumentali può anche essere facilitata dal tipo di stimolo usato nell'apprendimento pavloviano. Stimoli localizzati, ad esempio delle luci su una parete della gabbia che segnalano l'arrivo di cibo, possono elicitare fenomeni di *sign-tracking*, in cui il ratto apprende ad avvicinarsi alla luce. Così come il *goal-tracking* (avvicinamento al deposito), anche il *sign-tracking* può competere con la risposta strumentale e mascherare quindi l'effetto PIT. Per evitare questo problema, in molti studi sul PIT sono stati usati dei suoni. I suoni, in quanto più 'diffusi' e non altamente localizzabili come le luci, diminuiscono il rischio di *sign-tracking* (Holmes et al., 2010).

2.4.3 Interazioni con la svalutazione

In Corbit et al. (2007) è stata indagata la relazione tra PIT e stato motivazionale (fame/sazietà). Prima della fase di test, alcuni ratti sono stati saziati con del cibo: il risultato è stato che, oltre ad avere una baseline minore, questi ratti mostravano solo PIT specifico e non più PIT generale. Il PIT generale sembra quindi essere influenzato da questa procedura di svalutazione del rinforzo, mentre il PIT specifico ne è immune. Similmente, era già stato riportato come il PIT possa essere immune alla svalutazione in Rescorla (1994) ed in Holland (2004), ma in questi precedenti esperimenti non era stato usato il paradigma capace di distinguere tra PIT specifico e generale né erano state fatte lesioni che possano far luce sul tipo di PIT effettivamente elicitato.

Nel modello che andremo ad esporre, legheremo il PIT generale alla valutazione dell'utilità delle ricompense future (quindi suscettibile a svalutazione), mentre il PIT specifico sarà legato alle probabilità di successo dell'azione, ossia alle probabilità di ottenere la ricompensa collegata al premere la leva (indipendentemente dal suo valore), spiegando in questo modo il differente impatto della svalutazione sui due tipi di PIT.

2.4.4 Interazioni con le droghe d'abuso

Numerosi studi hanno mostrato come droghe d'abuso quali anfetamina, cocaina ed alcool possano avere particolari interazioni con il PIT (Wyvell & Berridge, 2001; LeBlanc et al., 2012; Corbit & Janak, 2007a). Sono stati mostrati sia effetti di 'potenziamento' del PIT, sia effetti di 'generalizzazione' del PIT.

Nel caso degli effetti di potenziamento, rilevati in esperimenti effettuati con una sola leva ed una sola ricompensa, è stato constatato che anfetamina e cocaina aumentano quantitativamente l'effetto PIT, ossia aumentano il numero aggiunto di pressioni della leva che si verificano durante la presenza dello stimolo condizionato rispetto alla *baseline* (Wyvell & Berridge, 2000; LeBlanc et al., 2012). Poiché sia anfetamina che cocaina aumentano il rilascio di dopamina nel *nucleus accumbens*, il loro effetto di potenziamento è in accordo con i dati di Lex & Hauber (2008) in cui era stata mostrata la relazione tra effetto PIT e recettori dopaminergici D1 e D2. È da notare comunque che in Wyvell & Berridge (2001) l'effetto di 'potenziamento' rispetto ai controlli si è verificato, a seguito di un trattamento di *sensibilizzazione*, anche in assenza di anfetamina durante il test. In questo caso quindi, il potenziamento non derivava dall'effetto dopaminergico immediato dell'assunzione di droga, ma da una sua azione a lungo termine sui circuiti neurali.

In altri esperimenti, usando due leve e due differenti ricompense, si è invece notato un effetto di 'generalizzazione'. Ad esempio in Shiflett (2012), i controlli *drug-free* mostravano un normale PIT specifico: ognuno dei due CS elicitava PIT solo in presenza della leva associata alla stessa ricompensa. I ratti soggetti ad un trattamento di *sensibilizzazione* all'anfetamina invece mostravano un effetto PIT verso entrambe le leve, a prescindere dal CS usato. L'effetto PIT era quindi diventato 'generalizzato', nel senso che non si applicava più solo alla leva corrispondente. Un simile risultato è stato ottenuto anche con l'alcool da Corbit & Janak (2007a), ma in questo caso l'effetto di generalizzazione si verificava solo per uno dei due CS, che era stato associato all'alcool.

Rispetto alla nostra ipotesi, questo effetto di *generalizzazione* potrebbe essere spiegato con la capacità delle droghe di elicitare un comportamento *habitual* (vedi ad es: Ostlund et al., 2010) e quindi elicitare un PIT generale anziché PIT specifico, così come ipotizzato per la differenza tra programmi di rinforzo RR e VI.

3 Ipotesi: tre principi del comportamento goal-directed

3.1 Spiegare il PIT

A prima vista, spiegare il PIT può sembrare semplice. I dati sui substrati neurali indicano chiaramente che esistono due tipi di PIT, specifico e generale. Il primo si verifica quando la ricompensa associata al CS è la stessa associata alla risposta strumentale. Il secondo, cioè il PIT generale, si verifica invece quando la ricompensa associata al CS è diversa da quella della risposta strumentale. Nel primo caso, potremmo pensare che il CS evoca la rappresentazione di una certa ricompensa e che questa a sua volta evoca la risposta collegata, aumentandone quindi la frequenza con cui viene eseguita. Nel secondo caso, il CS evoca la rappresentazione di una ricompensa che non è associata alla risposta strumentale ma ha comunque un effetto ‘motivazionale’ generale in quanto ricompensa, aumentando così il vigore (la frequenza) con cui vengono emesse le risposte, tra cui quella strumentale monitorata. Questa spiegazione è in effetti ciò che si trova in letteratura, ossia una spiegazione in termini di catene S-O-R (*stimulus-outcome-response*) per il PIT specifico, in cui lo stimolo condizionato (S) evoca il cibo (O) e questo a sua volta evoca la risposta (R); mentre per il PIT generale si ha un effetto motivazionale aspecifico che aumenta il vigore delle risposte. Nella letteratura del PIT, uno dei modelli più completi che comprende sia il condizionamento pavloviano che strumentale e che quindi può spiegare il PIT è probabilmente il modello associativo-cibernetico di Balleine e Dickinson (Balleine & Ostlund, 2007). In questo modello vi sono due modi tramite cui uno stimolo condizionato pavloviano, cioè uno stimolo associato ad un outcome (S-O) può influenzare le risposte strumentali. Uno dei modi è tramite l’associazione che si viene a creare tra il cibo inteso come stimolo e le risposte nella memoria S-R (vedi figura 3.1). Secondo Balleine & Ostlund (2007) poiché durante il condizionamento strumentale il cibo non solo segue la pressione della leva, ma precede anche la successiva pressione della leva, oltre ad apprendere il cibo come conseguenza della risposta (R-O

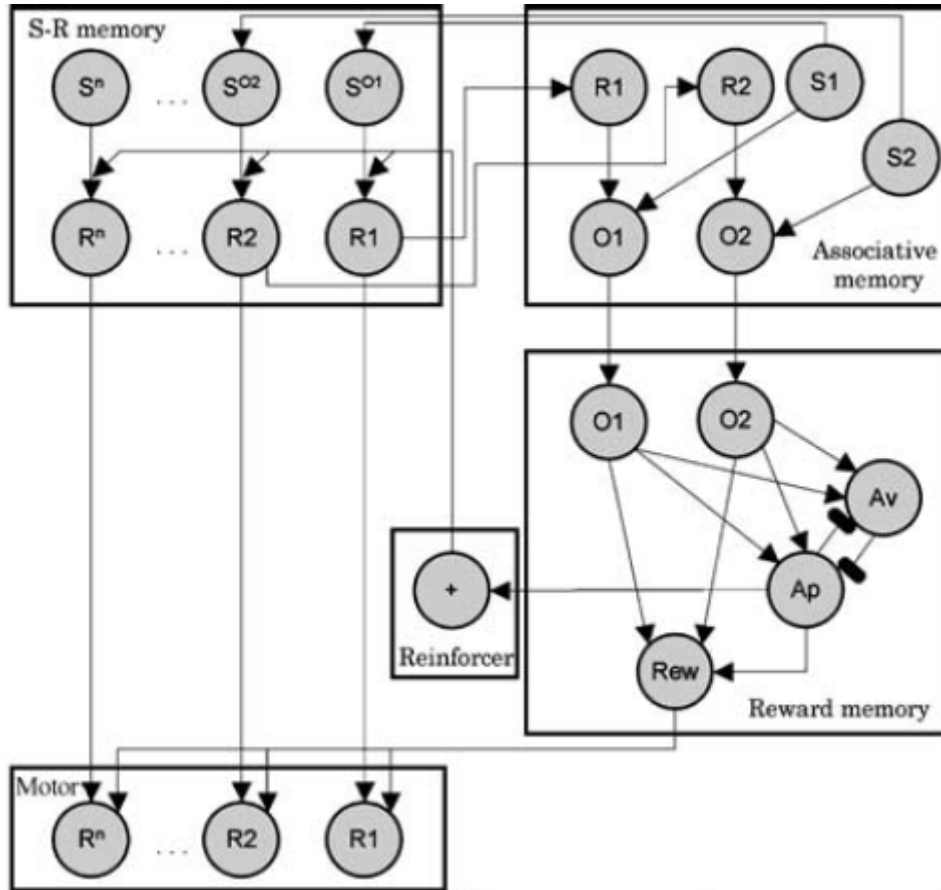


Figura 3.1: Il modello associativo-cibernetico di Balleine e Dickinson in Balleine & Ostlund (2007).

nella memoria associativa), si viene a formare anche un apprendimento di tipo S-R in cui al cibo *segue* la risposta (S^O-R nella memoria S-R). L'apprendimento pavloviano (S-O nella memoria associativa) può quindi evocare questo apprendimento strumentale S^O-R. Ad esempio, nella figura del modello, S1 nella memoria associativa è collegato a S^{O1} nella memoria S-R e quest'ultimo a sua volta è collegato ad R1. Si tratta all'atto pratico di una spiegazione in termini di catena S-O-R come enunciato precedentemente. Vi è però un secondo modo in cui gli stimoli pavloviani possono influenzare le risposte strumentali. Le associazioni S-O, possono tramite la memoria delle ricompense (*reward memory* in figura), attivare nel modello il nodo 'Rew' (un'aspettativa di valore) che a sua volta può potenziare le varie risposte strumentali, in maniera aspecifica (freccie che da Rew vanno a tutte le risposte). Questo secondo tipo di influenza potrebbe quindi spiegare il PIT generale, in cui uno stimolo condizionato associato ad una ricompensa può potenziare una risposta strumentale anche non associata allo stesso tipo di ricompensa,

tramite un'aspettativa generale di valore. Il fatto che il PIT generale sia collegato ad un'aspettativa di ricompensa spiegherebbe anche la sua capacità di essere influenzato dalle procedura di svalutazione, a differenza del PIT specifico che ne è immune.

Sebbene il modello associativo-cibernetico possa spiegare l'esistenza dei due tipi di PIT e la loro interazione con le procedure di svalutazione, alcuni aspetti importanti presenti nei dati rimangono senza risposta e suggeriscono che la spiegazione del PIT sia più complessa.

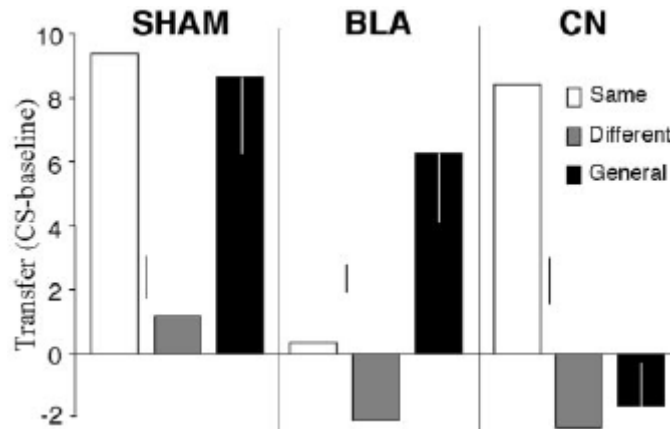


Figura 3.2: Dati da Corbit & Balleine (2005). Il gruppo di controllo senza lesioni (*sham*) mostra un effetto PIT sia nella condizione *same* (CS associato alla stessa ricompensa della risposta strumentale), sia nella condizione *general* (CS associato ad una ricompensa non usata come ricompensa strumentale). Da notare l'assenza di effetto PIT nella condizione *different*, ossia quando il CS è associato alla ricompensa dell'altra leva non disponibile durante il test. I gruppi BLA e CN mostrano l'effetto delle lesioni a carico dell'amigdala basolaterale (BLA) e dell'amigdala centrale (CN), che eliminano rispettivamente il solo PIT specifico o il solo PIT generale.

3.1.1 Assenza di PIT generale nella condizione di PIT specifico

Nella condizione di PIT specifico, un CS associato ad un certo cibo, potenzia una risposta strumentale diretta verso lo stesso cibo. Ci si potrebbe aspettare, secondo il modello associativo-cibernetico, che il CS possa in questa situazione elicitare PIT tramite entrambe le vie contemporaneamente, sia tramite la connessione $S1-S^{O1}-R1$, sia tramite $S1-O1-O1-Rew-R1$ (figura 3.1). Cioè ci si aspetterebbe che il CS possa avere sia un effetto di PIT specifico, in quanto evoca il cibo associato alla risposta strumentale, sia un effetto di PIT generale, in quanto evoca una aspettativa di valore capace di motivare in genere tutte le risposte strumentali. Dati sperimentali mostrano però che ciò non

accade. La figura 3.2 mostra come lesioni all'amigdala basolaterale (BLA), che fa parte del circuito del PIT specifico, eliminino totalmente l'effetto di PIT nella condizione 'same' (CS associato a stesso cibo della risposta). Viceversa, lesioni all'amigdala centrale (CN), che fa parte del circuito del PIT generale, non hanno effetti sulla condizione 'same'. Se ne deduce quindi che in questa condizione è espresso solo il PIT specifico, e non il PIT generale. Eventualmente, sarebbe possibile spiegare questo effetto all'interno del modello associativo-cibernetico assumendo che la rappresentazione del cibo O1 nella memoria associativa (vedi figura 3.1) abbia un valore massimo di attivazione, e che questo valore massimo sia già raggiunto tramite il contributo dell'associazione R1-O1, per cui il contributo di S1 (il CS pavloviano) nell'attivare 'la via del PIT generale' sarebbe nullo. L'associazione R1-O1 è però a sua volta attivata da R1 nella memoria S-R. Quindi la saturazione di O1 nella memoria associativa sarebbe dipendente dall'attivazione di R1 nella memoria S-R. Da questo conseguirebbe che, all'aumentare della forza delle associazioni S-R (e quindi di R1 nella memoria S-R), aumenta la probabilità di saturare O1 nella memoria associativa e quindi di rendere nullo il contributo della 'via del PIT generale'. Eppure, in esperimenti di PIT con una sola leva e con programmi ad intervallo variabile, in cui il comportamento era probabilmente di tipo *habitual* e quindi con una associazione S-R forte, sono stati rilevati effetti PIT dipendenti dalla via del PIT generale e non dalla via del PIT specifico. Al contrario ci si sarebbe aspettati che con una componente S-R forte, R1 in S-R dovrebbe attivare (e saturare) l'associazione R-O in memoria associativa. Eventualmente, questa spiegazione potrebbe essere salvata introducendo altre assunzioni sulla forza dell'associazione R-O o variando opportunamente la soglia massima di attivazione di O1. In ogni caso, tutte questi dettagli costituirebbero delle aggiunte al modello associativo-cibernetico, che nella sua versione base non sembra spiegare l'assenza di PIT generale in presenza di PIT specifico.

3.1.2 Assenza di PIT generale verso alcune risposte strumentali

Vi è un secondo aspetto per cui la spiegazione del modello associativo cibernetico è incompleta. Nei dati mostrati in figura 3.2 possiamo notare come nella condizione *different* non vi è effetto PIT (né specifico né generale). La condizione *different* corrisponde al caso in cui il CS è associato alla ricompensa della risposta strumentale non disponibile, ossia alla ricompensa della leva non presente nella fase di test. Ci si aspetterebbe che il CS, non essendo lo stesso della risposta strumentale disponibile, non riesca ad elicitarne PIT specifico ma elicitare comunque PIT generale, tramite un effetto motivazionale aspecifico sulla risposta strumentale disponibile (effetto di Rew su tutte le risposte motorie

nel modello associativo-cibernetico). Al contrario, non vi è alcun effetto PIT visibile, la risposta strumentale rimane ai livelli di baseline ed anche effettuando lesioni a carico del circuito del PIT specifico o del PIT generale, il livello rimane attorno alla baseline. Il modello associativo-cibernetico non offre spiegazioni in merito: non vi è nel modello una connessione che possa spiegare come mai l'effetto aspecifico della 'via del PIT generale' possa essere soppresso verso alcune risposte strumentali. Anche gli articoli che contengono i dati di questa 'assenza del PIT generale' si limitano a suggerire un effetto inibitorio non meglio specificato (Corbit & Balleine, 2005) o a suggerire a livello neurale dove potrebbe verificarsi questo effetto inibitorio (Corbit & Balleine, 2011), senza però spiegare esattamente la motivazione della sua presenza. Vi è quindi un effetto inibitorio capace di sopprimere il PIT generale in attesa di essere spiegato. Questo effetto non è sempre della stessa entità, in alcuni casi è possibile avere un effetto di PIT 'parziale' nella condizione *different*, come ad esempio in figura 3.3.

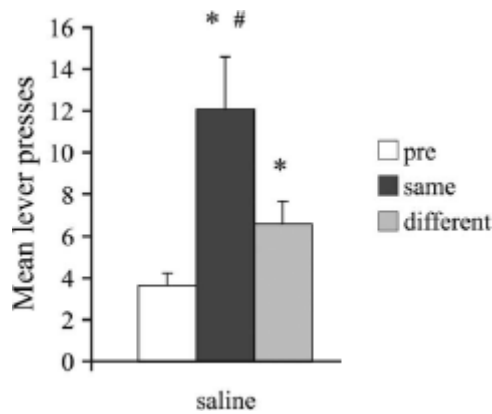


Figura 3.3: Dati da Corbit & Janak (2007b). Anche nella condizione *different*, in cui il CS è associato alla ricompensa della leva non presente è a volte possibile ottenere un effetto PIT, ossia una risposta maggiore rispetto alla condizione *pre-CS* (*baseline*).

3.1.3 Assenza di PIT specifico nei paradigmi con una sola leva

La figura 3.4 mostra i risultati di Hall et al. (2001), in cui è stato usato un singolo CS ed una sola risposta strumentale, associati entrambi alla stessa ricompensa. Il programma di rinforzo era inoltre di tipo ad intervallo variabile. I dati mostrano come solo il gruppo di controllo ed il gruppo con lesione allo *shell* mostrano un effetto PIT, ossia un innalzamento rispetto alla *baseline* (ITI) in presenza del CS. Nonostante sia stata usata la stessa ricompensa per CS e risposta strumentale, è la lesione al *core* (circuito PIT generale) ad impedire l'effetto PIT e non quella allo *shell* (circuito PIT specifico). All'interno del

3 Ipotesi: tre principi del comportamento *goal-directed*

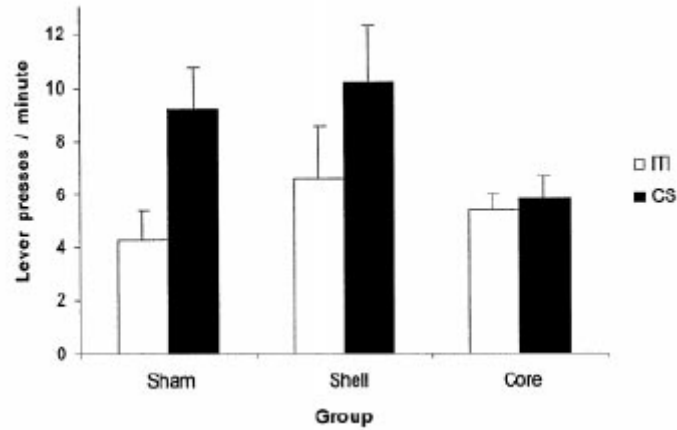


Figura 3.4: Dati da Hall et al. (2001). Il CS aumenta la risposta rispetto alla baseline (ITI) solo nel gruppo con lesione allo *shell* e nel gruppo di controllo. Nonostante sia stata usata la stessa ricompensa per CS e risposta strumentale, è la lesione al *core* (circuito PIT generale) ad impedire l'effetto PIT e non quella allo *shell* (circuito PIT specifico).

modello associativo-cibernetico, si potrebbe argomentare che nel caso della procedura ad intervallo, poiché è necessario attendere del tempo tra una pressione di una leva e l'altra, l'associazione tra cibo e successiva pressione della leva sia più debole (associazione S^{O1} -R1). Questo implicherebbe che il PIT specifico è più debole (o addirittura così debole da non essere rilevabile) nella procedura ad intervallo usata da Hall et al. (2001) e da Holland & Gallagher (2003). Non spiegherebbe però come mai usando invece la procedura ad intervallo con due leve, come nel caso di Blundell et al. (2001), si riesca invece ad ottenere un PIT di tipo specifico, eliminabile tramite lesioni all'amigdala basolaterale. Come già anticipato nel paragrafo 2.4.1, la nostra idea è che la differenza tra usare una procedura ad intervallo con una sola leva o con due leve, sia nel fatto che nel primo caso è favorita la formazione di un comportamento strumentale di tipo *habitual*, mentre con l'uso delle due leve il comportamento rimane *goal-directed*. Poiché riteniamo che il PIT specifico sia legato alla valutazione della probabilità di ottenere un certo fine tramite l'azione, ne consegue chiaramente che se l'azione è di tipo stimolo-risposta, cioè *habitual*, non essendoci valutazione delle conseguenze dell'azione non vi è nemmeno PIT specifico.

3.2 Tre principi: contesto, efficacia, utilità

I dati relativi alla fase di test del PIT ci pongono di fronte a tre condizioni con tre tipi di effetti: il PIT specifico, il PIT generale e l'assenza del PIT generale quando il CS è associato alla ricompensa della leva non presente. Esponiamo qui la nostra ipotesi secondo cui questi tre tipi di effetti sono funzionalmente legati a tre aspetti della valutazione dell'azione nel caso di comportamento *goal-directed*.

Nei paragrafi precedenti abbiamo inoltre segnalato ulteriori due effetti: l'assenza di PIT generale nella condizione di PIT specifico e viceversa la presenza di PIT generale al posto del PIT specifico nel caso di comportamento *habitual*. Per quanto riguarda l'assenza di PIT generale nella condizione di PIT specifico, sarà chiaro nella trattazione seguente che considerare il PIT generale come dovuto alla presenza di una ricompensa 'aggiuntiva' rispetto a quella strumentale, ne elimina automaticamente la presenza nella condizione di PIT specifico (in quanto il CS presente non segnala una ricompensa diversa da quella strumentale)¹. Per quanto riguarda invece il PIT nel comportamento *habitual*, l'ipotesi funzionale che descriveremo è limitata al comportamento *goal-directed* e non tratta quindi direttamente la questione della presenza di PIT generale al posto di PIT specifico nel caso di comportamento *habitual*. È però facile ipotizzare che se la nostra ipotesi fosse corretta, il PIT specifico, riferendosi alla probabilità di successo dell'azione non potrebbe avvenire nel caso di comportamento *habitual*, in cui non vengono considerate le conseguenze dell'azione, mentre il PIT generale, riferendosi più generalmente alla presenza motivante di ricompense, potrebbe essere ancora presente, anche al posto del PIT specifico.

Come già ricordato, un'azione strumentale, quale il premere una leva che conduce al cibo, può avvenire sia in forma di risposta *habitual*, sia in forma *goal-directed*. Nel caso *habitual* la risposta è di tipo S-R (stimolo-risposta) - è una sorta di 'reazione' ad uno stimolo - e non vi è una valutazione delle conseguenze dell'azione. Nel caso *goal-directed*, la risposta è considerata di tipo A-O (*action-outcome*) ossia legata alla valutazione delle sue conseguenze. A livello comportamentale, è possibile distinguere i due tipi di comportamento strumentale tramite i paradigmi della 'svalutazione' e della 'degradazione della contingenza'. Nel paradigma di svalutazione viene svalutato il valore del cibo (ad esempio tramite sazietà) mentre nel paradigma della degradazione della contingenza vie-

¹Non è escluso che a livello neurale il PIT generale nella condizione di PIT specifico venga inibito tramite lo stesso meccanismo usato per l'assenza di PIT generale quando la ricompensa del CS è associata alla leva non presente. Anche a livello funzionale potrebbe esserci una spiegazione comune ma non indagheremo qui questa ipotesi.

ne alterata la capacità della leva di fornire il cibo (es: premere la leva anziché fornire il cibo, impedisce che il cibo arrivi). Nel caso del comportamento *habitual*, il ratto (o altro soggetto) continuerà a premere la leva, mentre nel caso del comportamento *goal-directed* smetterà di premerla in quanto sarà cambiata la capacità dell'azione di fornire un certo cibo (degradazione della contingenza) o il valore del cibo ottenibile (svalutazione).

Il comportamento *goal-directed* si distingue dunque per il suo valutare le conseguenze dell'azione. La nostra ipotesi è che i vari aspetti del PIT, nel comportamento *goal-directed*, siano riconducibili a tre aspetti di questa valutazione:

Contesto: un'azione diretta ad un fine deve essere eseguita nel giusto contesto (inutile premere l'interruttore se è andata via la luce)

Efficacia: un'azione può essere più o meno efficace nel raggiungere un fine (es: comprare un biglietto della lotteria ha poche probabilità di vittoria)

Utilità: il fine dell'azione può essere più o meno utile (es: ottenere cibo è utile se sono affamato, meno, o per nulla, se sono sazio)

Il PIT generale è riconducibile al principio di *utilità*: un CS associato a del cibo, evoca una ricompensa nell'immediato futuro. La nostra ipotesi è che all'atto di valutare le conseguenze future dell'azione, questa ricompensa venga aggiunta nello scenario, aumentando così la motivazione ad effettuare l'azione. Se però si è sazi, questa ricompensa aggiuntiva chiaramente non ha valore e l'effetto motivazionale sull'azione scompare. È stato infatti dimostrato che la sazietà elimina il PIT generale (Corbit et al., 2007). Inoltre, l'effetto motivazionale del PIT generale è possibile solo se il CS evoca una ricompensa aggiuntiva rispetto a quella già prevista dall'azione. In accordo con ciò, i dati ottenuti tramite programmi di rinforzo *random-ratio*, in cui il comportamento tende ad essere *goal-directed*, mostrano come nella condizione di PIT specifico (il CS è associato alla stessa ricompensa della leva) non si verifica il PIT generale (Corbit & Balleine, 2005, 2011).

Il PIT specifico è invece riconducibile al principio di *efficacia*. Il CS associato alla stessa ricompensa della leva predice che questa ricompensa sarà presente nel futuro. All'atto della valutazione dell'azione, il CS funziona quindi da segnale che indica una maggiore probabilità di ottenere la ricompensa associata all'azione. Poiché nel PIT specifico si tratta di un *aumento della probabilità* di ottenere il cibo, questo effetto è immune alla svalutazione: che io sia sazio o meno, comunque il CS predice una maggior probabilità di ricompensa, quindi rispetto all'assenza del CS, la risposta sarà comunque

valutata maggiormente. A conferma del legame tra PIT specifico e probabilità, in uno studio PIT con soggetti umani, Trick et al. (2011) hanno rilevato come CS più predittivi (cioè con maggiore probabilità di ricompensa) inducevano un effetto PIT più forte.

Infine, l'assenza dell'effetto di PIT generale quando il CS è associato alla ricompensa della leva non presente è riconducibile al principio del *contesto*. Ipotizziamo infatti che la presenza della leva agisca come stimolo discriminante in maniera inibitoria rispetto alla ricompensa della leva non presente. Ossia la presenza della leva segnala non solo che è possibile ottenere una certa ricompensa, ma anche che *non* è possibile ottenere la ricompensa dell'altra leva non presente. L'effetto di PIT generale sarebbe quindi inibito dalla presenza della leva associata alla ricompensa diversa.

In altre parole, il PIT specifico rappresenta la capacità di un agente di sfruttare i segnali che indicano che la presenza di una certa ricompensa è più probabile nell'ambiente rispetto a quando il segnale è assente e che quindi un'azione diretta ad ottenerla ha maggiori probabilità di successo (*efficacia*). Il PIT generale rappresenta invece l'abilità di usare segnali che indicano la presenza di altre ricompense 'aggiuntive' nell'ambiente e che quindi motivano ad agire in quanto costituiscono un valore aggiunto (*utilità*). L'inibizione del PIT generale rappresenta invece la capacità di tenere conto del contesto e di inibire segnali di ricompense che si sanno non ottenibili al momento tramite l'azione (*contesto*).

Abbiamo voluto quindi esprimere questa ipotesi funzionale sui vari effetti PIT tramite un modello computazionale. In attesa di poter sviluppare in futuro un modello che comprenda anche i sottostanti meccanismi neurali, abbiamo quindi iniziato a lavorare ad un modello puramente funzionale espresso tramite una *sigmoid belief network*, ossia un tipo di rete bayesiana. Le reti bayesiane sono già state applicate sia al condizionamento pavloviano (Courville et al., 2004, 2005, 2006; Gershman & Niv, 2012) sia al condizionamento strumentale (Solway & Botvinick, 2012; Pezzulo et al., 2013). Il nostro modello prova ad estendere uno dei lavori di Courville in ambito pavloviano al caso strumentale, al fine di simulare gli effetti di interazione tra i due condizionamenti, ossia il PIT. Nel prossimo capitolo, descriveremo quindi l'approccio bayesiano in generale ed il lavoro di Courville et al. (2005) in particolare.

4 Reti bayesiane e modelli con variabili latenti

4.1 Modelli probabilistici

Per costruire il nostro modello funzionale del PIT abbiamo deciso di usare un approccio di tipo ‘bayesiano’, ossia probabilistico. L’uso del termine ‘bayesiano’ deriva dal fatto che nell’approccio probabilistico assume un ruolo centrale il teorema di Bayes.

Il teorema di Bayes è il seguente:

$$p(h|d) = \frac{p(d|h)p(h)}{p(d)} \quad (4.1)$$

ossia, la probabilità di un certo evento h dato un insieme di osservazioni d ($p(h|d)$) è pari alla probabilità che si verifichino le osservazioni d dato l’evento h moltiplicata per la probabilità *a priori* di h e diviso la probabilità *a priori* che si verifichino le osservazioni d . Il teorema fondamentale indica come tenere conto di un insieme di osservazioni d nel calcolare il valore di probabilità dell’evento h . In altre parole, trasforma la probabilità *a priori* di h in una probabilità *a posteriori* di h che tiene conto dell’insieme di osservazioni d . Il teorema di Bayes indica come ‘aggiornare’ la nostra credenza sul verificarsi di h in maniera ottimale rispetto ai dati d a disposizione. La ‘nuova credenza’ in merito ad h , aggiornata con i dati d , ossia la distribuzione *a posteriori* $p(h|d)$ può essere usata a sua volta come una distribuzione *a priori* per integrare un nuovo set di osservazioni z , semplicemente applicando nuovamente il teorema di Bayes (4.1).

L’approccio bayesiano viene usato, sempre più spesso, in vari campi, dalla chimica (Hibbert & Armstrong, 2009) all’astrofisica (Loredò, 1992), dall’economia (Poirier, 2006) alla genetica (Beaumont & Rannala, 2004) ed in numerosi altri campi ancora. Nonostante il teorema di Bayes abbia in realtà più di due secoli, la sua applicazione è stata per lungo tempo ristretta a casi particolari con poche variabili e con particolari distribuzioni di probabilità in cui era effettivamente possibile effettuare i calcoli neces-

sari. L'invenzione di nuovi metodi di approssimazione per il calcolo delle distribuzioni di probabilità (es: i metodi Monte Carlo Markov Chain) nonché la disponibilità di computer sempre più potenti e di software specializzati ha permesso quella che da alcuni è stata chiamata 'la rivoluzione bayesiana' (Beaumont & Rannala, 2004).

Tra i campi investiti da questa 'rivoluzione' vi sono anche le scienze cognitive: *Trends in Cognitive Science* ha pubblicato nel 2006 uno *special issue* sui modelli probabilistici della cognizione (Chater et al., 2006). Per quanto riguarda il tema di questa tesi, l'approccio probabilistico è stato utilizzato per creare dei modelli, tramite reti bayesiane, sia del condizionamento pavloviano (Courville et al., 2004) che del condizionamento strumentale (Solway & Botvinick, 2012).

4.2 Reti bayesiane

Le reti bayesiane sono un tipo di modello grafico che permette di rappresentare delle relazioni di dipendenza (o indipendenza) tra variabili¹. Ad esempio, nella rete in figura 4.1 le due variabili a e b , rappresentate come nodi di un grafo, sono *a priori* indipendenti mentre c dipende sia da a che da b .

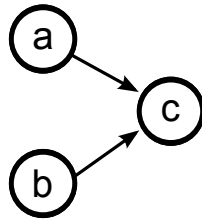


Figura 4.1: Una semplice rete bayesiana con 3 variabili a, b, c . Le reti bayesiane sono un grafo in cui i nodi rappresentano le variabili e gli archi permettono di esprimere le dipendenze tra di esse.

La rete in figura 4.1 stabilisce quindi che possiamo esprimere la distribuzione di probabilità congiunta delle tre variabili $p(a, b, c)$ nel seguente modo:

$$p(a, b, c) = p(a)p(b)p(c|a, b) \quad (4.2)$$

ossia la probabilità congiunta di a, b, c equivale alle probabilità indipendenti di a e b moltiplicate per la probabilità di c dati a e b . Le reti bayesiane quindi esprimono in un

¹Per la precisione, il grafo di una rete bayesiana in sé stabilisce le indipendenze, l'effettiva presenza delle dipendenze (o di ulteriori indipendenze) dipende dai particolari valori delle distribuzioni di probabilità delle variabili 'dietro' il modello grafico.

certo senso ‘come le distribuzioni di probabilità delle varie variabili sono collegate tra di loro’. Ossia indicano quali inferenze possiamo fare a partire da certe variabili osservate su quelle non osservate. Per fare un esempio in tema, la figura 4.2 mostra il modello di Solway & Botvinick (2012) con cui viene descritto il comportamento *goal-directed*. Nel

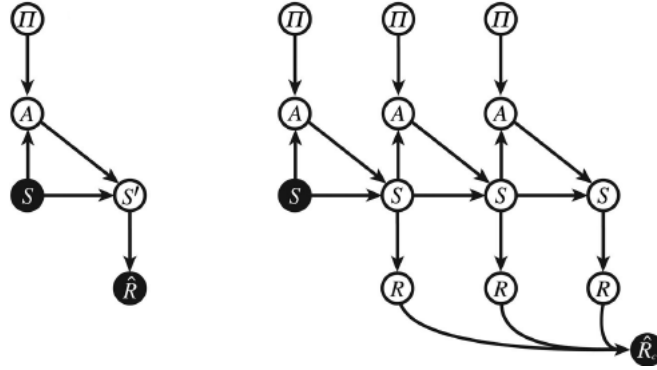


Figura 4.2: Da Solway & Botvinick (2012).

modello, la probabilità di ottenere una ricompensa \hat{R} dipende dallo stato futuro S' - questo a sua volta dipende dallo stato presente S e dall'azione A compiuta². Data una certa situazione S e la volontà di ottenere una certa ricompensa \hat{R} (nodi ombreggiati, cioè ‘osservati’) è possibile utilizzare la rete per inferire quale sia l'azione A ottimale da compiere, ossia l'azione che più probabilmente porta alla ricompensa \hat{R} . Nel lato destro della figura, il modello viene ‘replicato’ più volte per prendere in considerazione una serie di tre scelte di azioni consecutive, ognuna delle quali conduce a partire da uno stato S di partenza ad uno nuovo stato S ; ognuno di questi stati può procurare una certa ricompensa R e queste ricompense a loro volta sono ‘sommate’ in una ricompensa totale \hat{R} . Solway & Botvinick (2012) mostrano come questo modello possa catturare il comportamento di scelta di un ratto all'interno di un labirinto in cui una serie di bivi lo separano dal cibo. Inoltre riscontrano possibili parallelismi tra l'attività di alcuni neuroni e le probabilità calcolate durante le inferenze effettuate tramite il modello. Nel prossimo paragrafo vedremo invece un esempio di modello bayesiano del condizionamento pavloviano.

²Nel modello A dipende a sua volta dallo stato presente S e dalla *policy* Π

4.3 Apprendimento pavloviano e cause latenti: il modello di Courville

L'apprendimento pavloviano viene solitamente pensato in termini di associazioni S-O, ossia di associazioni tra uno stimolo S (es: un suono) ed un *outcome* O (es: del cibo). Queste associazioni si formerebbero in quanto il cibo O è presentato in maniera contingente al suono S. Modelli associazionisti come il modello di Rescorla-Wagner (Rescorla & Wagner, 1972) o Pearce (Pearce, 1994) mostrano come l'associazione venga appresa, ossia come cambi la forza dell'associazione in funzione dei trial pavloviani vissuti, ossia della contingenza sperimentata tra S ed O. In particolare, l'associazione S-O in questi modelli si focalizza sulla capacità dello stimolo condizionato S di predire la presenza del rinforzo O. Nel tempo, le teorie associazioniste hanno dato conto di numerosi fenomeni che si verificano durante l'apprendimento pavloviano. I modelli associazionisti S-O non sono però l'unica spiegazione possibile dell'apprendimento pavloviano. Ad esempio, in una serie di articoli, Courville e colleghi (Courville et al., 2004, 2005, 2006) mostrano come sia possibile utilizzare modelli bayesiani con variabili latenti per spiegare il condizionamento pavloviano. L'utilizzo di questi modelli permette non solo una 'spiegazione alternativa' del condizionamento pavloviano, ma anche di spiegare effetti non spiegati precedentemente (Courville et al., 2004) o non riconciliabili all'interno di uno stesso modello associazionista. Ad esempio, in Courville et al. (2005), tramite il modello bayesiano con cause latenti vengono modellati sia l'effetto *overshadowing* che l'effetto *summation*, fenomeni che in genere non erano predicibili all'interno dello stesso modello, ad esempio usando il solo modello di Rescorla-Wagner o il solo modello di Pearce.

Nel modello di Courville et al. (2005), essendo un approccio di tipo generativo anziché discriminativo, l'apprendimento pavloviano non è limitato al solo apprendere la relazione predittiva tra S ed O³. Al contrario, secondo Courville l'apprendimento pavloviano implica apprendere quali siano le cause sottostanti la correlazione tra S ed O (variabili latenti). In altre parole, nel modello di Courville il ratto apprende non semplicemente che S è predittivo di O, ma apprende che esiste una variabile latente, una causa nascosta che causa sia S che O (vedi figura 4.3). In effetti, in un tipico esperimento pavloviano, non è il suono S a 'causare' il cibo O, ma è lo sperimentatore (variabile latente, 'nascosta') a causare l'apparizione di entrambi S ed O.

³Negli approcci generativi viene modellata la probabilità congiunta delle variabili osservate, mentre negli approcci discriminativi viene modellata solo la probabilità di alcune variabili condizionata alle altre. Ad esempio, nei modelli pavloviani discriminativi viene solitamente modellata solo la probabilità di O dato S, ossia $p(O|S)$.

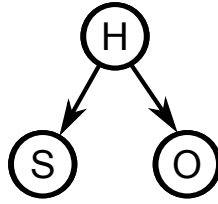


Figura 4.3: Il condizionamento classico può essere pensato, anziché in termini di associazioni S-O, come l'apprendimento delle variabili latenti che causano sia S che O, ossia ne spiegano la correlazione.

Come nei modelli associazionisti, vedere S predice O, in quanto se il ratto vede S può dedurre che la causa latente H è attiva e che quindi presto accadrà O. Il ratto però non apprende solo ad inferire la probabilità di O in presenza di S (cioè ad inferire $p(O|S)$). Apprendendo l'intero modello 'S-H-O', il ratto apprende anche la probabilità di S quando H è attiva $p(S|H)$ e la probabilità a priori di H $p(H)$. Il ratto può quindi anche predire quanto spesso accadrà S. In questo senso il modello non è solo discriminativo ma è generativo ossia permette di fare inferenze su entrambe le variabili osservabili (S ed O); a differenza dei modelli associazionisti in cui il ratto non impara a predire S, ma solo O in presenza di S.

Nel modello di Courville et al. (2005), il ratto durante l'apprendimento pavloviano usa l'inferenza per determinare la struttura stessa del modello generativo che determina le sue osservazioni. Ossia tenta di inferire quante sono le cause nascoste in gioco e quali variabili osservabili siano sotto l'influenza di ciascuna causa. Courville et al. (2005) mostra come nei paradigmi pavloviani che prevedono più stimoli condizionati il numero di cause nascoste ritenuto più probabile ed i loro collegamenti con gli osservabili possano variare in funzione dell'esperienza (vedi figura 4.4). La flessibilità della struttura del modello,

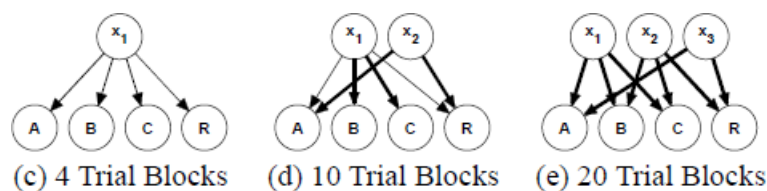


Figura 4.4: Da Courville et al. (2005). Durante l'apprendimento pavloviano, il ratto tenta di stabilire quali e quante siano le cause latenti (x_1, x_2, x_3) che determinano le correlazioni tra i vari stimoli condizionati A, B, C ed il rinforzo R. La figura mostra come, con il progredire dei trial, il modello di cause nascoste ritenuto più probabile passi da una singola causa a tre cause distinte.

ossia la sua capacità di inferire la presenza di diverse cause a seconda dell'esperienza, fa

in modo che il modello possa spiegare più agevolmente dei precedenti modelli associativi i fenomeni pavloviani che avvengono in presenza di stimoli condizionati multipli. Poiché la stessa struttura del modello è inferita tramite il teorema di Bayes, l'approccio risulta caratterizzato da un aspetto normativo altrimenti assente nei modelli associazionisti. Ossia, la scelta di 'legare' alcuni stimoli condizionati tra loro tramite una stessa causa latente (vedi figura 4.4) dipende dall'applicazione del teorema di Bayes all'esperienza dei trial. Nei precedenti modelli associativi, la scelta di legare più stimoli assieme in una 'configurazione di stimoli' era invece effettuata tramite delle scelte *ad hoc* (Courville et al., 2005).

5 Modello del PIT

5.1 Struttura del modello

Per costruire il modello del PIT, ci siamo ispirati al modello di Courville et al. (2005) del condizionamento pavloviano descritto nel capitolo precedente. Abbiamo ‘esteso’ il modello al comportamento strumentale nel modo più semplice possibile, ossia aggiungendo un singolo nodo osservabile A che rappresenta l’azione di premere la leva. Tale nodo non ha genitori, in quanto non dipende da nessuno degli altri nodi: né dagli osservabili (suoni, cibi, leve), né dalle cause latenti.

Il nostro modello è quindi una rete bayesiana (vedi figura 5.1) composta dai seguenti tipi di nodi:

- Osservabili: nodi che rappresentano oggetti quali leve ($L1, L2, L3$), suoni ($S1, S2, S3$), cibi ($F1, F2, F3$)
- Cause nascoste: nodi che rappresentano cause ipotetiche ($H1, H2, H3, H4, H5$) che possono spiegare l’osservazione (o la mancata osservazione) degli oggetti nel mondo
- Azioni: un singolo nodo che rappresenta l’azione del premere una leva (A) e che può influenzare la presenza o meno di certi oggetti nel mondo

Rispetto al modello di Courville, in questo modello il ratto non impara solo quali siano i collegamenti tra variabili latenti ed osservabili, ma anche i collegamenti tra l’azione e gli osservabili. Impara cioè, non solo quali siano le ‘cause nascoste’ che agiscono nel mondo, ma anche come la sua azione di premere le leve influenzi il mondo. In particolare, nel caso del PIT, il ratto (o altro agente) dovrà imparare che l’azione premere la leva ha un influenza positiva (cioè un ‘peso’ positivo nel modello) sulla presenza o meno di alcuni cibi. Abbiamo collegato il nodo A ai soli cibi $F1$ ed $F2$, ossia alle ricompense strumentali delle leve $L1$ ed $L2$. Solo i cibi $F1$ ed $F2$ infatti possono accadere *dopo* l’azione di premere la leva, mentre gli altri osservabili o sono già presenti nel contesto (es: la leva) oppure

non sono presenti, nemmeno dopo la pressione della leva (es: il cibo F3, non usato nelle sessioni strumentali). Questa scelta di collegare l'azione A ai soli osservabili F1 ed F2, sottintende quindi la capacità da parte dell'agente di intendere la nozione di *causalità*, dove i possibili effetti di una propria azione sono necessariamente *successivi* all'azione.

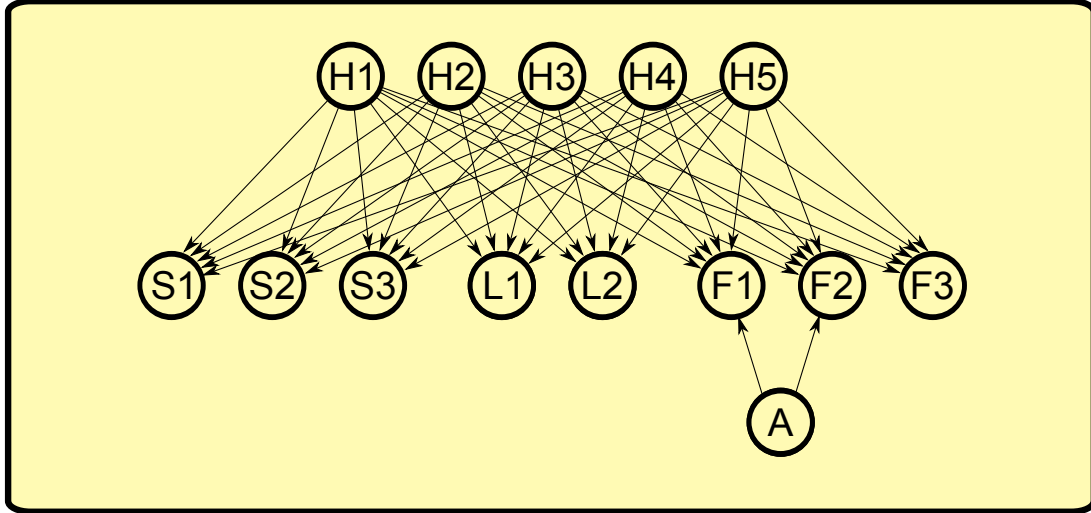


Figura 5.1: La rete bayesiana per simulare il PIT. Sono presenti tre tipi di nodi: oggetti osservabili nel mondo quali leve (L1, L2), suoni (S1, S2, S3) e cibi (F1, F2, F3), cause nascoste che influenzano la presenza degli oggetti nel mondo (H1, H2, H3, H4, H5) e l'azione di premere la leva (A). L'azione di premere la leva è collegata solo con i cibi F1 ed F2 in quanto può influenzare solo la loro presenza e non quella degli altri osservabili.

La rete, come nel modello di Courville et al. (2005), è implementata tramite una *sigmoid belief network*. Ogni nodo può assumere un valore compreso tra 0 ed 1, che rappresenta la probabilità dell'evento rappresentato dal nodo di accadere. Un valore di 0.75 per il nodo L1 indicherà quindi che abbiamo il 75% di probabilità di osservare la leva L1. L'attivazione dei nodi dipende dalla seguente formula:

$$P(S_i|w_i, x) = (1 + \exp(-(w_i)^T x - w_{bias}))^{-1} \quad (5.1)$$

dove x è il vettore delle attivazioni dei nodi genitori (le cause latenti ed il nodo azione) e w_i è il vettore dei pesi degli archi dai genitori verso il nodo S_i . Dati x e w_i la probabilità di ogni nodo S_i di essere attivo è pari alla funzione sigmoide della somma delle attivazioni x dei suoi genitori ognuna moltiplicata per il rispettivo peso dell'arco diretto al nodo (w_i), più un peso di *bias* w_{bias} . In altre parole, ogni peso rappresenta l'influenza della presenza di un nodo sui nodi figli. Questa influenza può essere positiva o negativa, aumentando

o diminuendo quindi le probabilità che il nodo figlio si verifichi. La funzione sigmoide assicura che il valore delle attivazioni dei nodi figli sia sempre compreso tra 0 ed 1. Infine, il peso di *bias* rappresenta la probabilità di ogni nodo di accadere (essere attivo) quando i suoi genitori non sono attivi.

5.2 Apprendimento del modello

In Courville et al. (2005), la struttura del modello è appresa, ossia vengono appresi sia il numero di cause nascoste presenti, sia il numero di archi tra le varie cause latenti e gli osservabili. Per semplicità, nella versione attuale del nostro modello, la struttura è fissa, ossia vi sono sempre cinque cause latenti collegate a tutti gli osservabili. Vengono quindi appresi solo i pesi degli archi. Imparare i pesi degli archi significa che il modello deve scoprire come le cause nascoste (e l'azione di premere la leva) stiano influenzando le probabilità di verificarsi di ogni osservabile.

5.2.1 Dati

I dati usati per far apprendere i pesi sono dei dati artificiali che rappresentano le fasi del paradigma PIT. I dati sono costituiti da vettori di 0 ed 1 che rappresentano le osservazioni all'interno di un singolo 'trial' pavloviano o strumentale¹. Nel caso dell'apprendimento pavloviano tra un suono S1 ed il cibo F1, avremo un vettore in cui S1 ed F1 sono valorizzati ad 1 (presenti) mentre tutti gli altri osservabili (es: le leve L1 ed L2) hanno valore 0 in quanto non presenti. Poiché nell'apprendimento pavloviano il suono tipicamente ha una lunga durata (es: 120 secondi) all'interno della quale il cibo viene presentato ad intervalli, abbiamo introdotto per l'apprendimento pavloviano anche dei trial in cui solo il suono è attivo, mentre il cibo è assente. In modo simile sono stati costruiti i dati per l'apprendimento strumentale, in cui abbiamo tre possibili situazioni:

- leva presente, azione non effettuata, cibo assente
- leva presente, azione effettuata, cibo presente
- leva presente, azione effettuata, cibo assente (la pressione della leva non è rinforzata in maniera continua)

¹In realtà rappresentare l'esperienza del condizionamento pavloviano e soprattutto del condizionamento strumentale come una serie di 'trial' è una semplificazione. Per un'interessante discussione di alcuni problemi di questa semplificazione si rimanda a *'The trial problem'* in Gallistel & Gibbon (2000).

La tabella 5.1 riassume tutte le possibili combinazioni usate. Ogni vettore di osservazioni è stato replicato 30 volte, per un totale di 180 trial pavloviani e 180 trial strumentali.

F1	F2	F3	L1	L2	S1	S2	S3	A	Tipo di trial
1	0	0	0	0	1	0	0	0	Pavloviano
0	0	0	0	0	1	0	0	0	Pavloviano
0	1	0	0	0	0	1	0	0	Pavloviano
0	0	0	0	0	0	1	0	0	Pavloviano
0	0	1	0	0	0	0	1	0	Pavloviano
0	0	0	0	0	0	0	1	0	Pavloviano
1	0	0	1	0	0	0	0	1	Strumentale
0	0	0	1	0	0	0	0	1	Strumentale
0	0	0	1	0	0	0	0	0	Strumentale
0	1	0	0	1	0	0	0	1	Strumentale
0	0	0	0	1	0	0	0	1	Strumentale
0	0	0	0	1	0	0	0	0	Strumentale

Tabella 5.1: Vettori delle osservazioni usati per l'apprendimento

5.2.2 Monte Carlo Markov Chain

L'apprendimento dei dati consiste nel trovare, tramite inferenza bayesiana, le combinazioni di pesi nel modello che hanno maggiori probabilità di spiegare i dati osservati. Più formalmente significa che, a partire da una distribuzione *a priori* dei pesi, viene trovata una distribuzione di pesi *a posteriori* in funzione dei dati osservati.

L'inferenza è stata effettuata tramite il metodo Monte Carlo Markov Chain (MCMC), usando il software WinBUGS (Lunn et al., 2000). L'utilizzo di tale software ci ha permesso di delegare al programma stesso la scelta e l'ottimizzazione degli algoritmi di *sampling* del metodo MCMC. WinBUGS non offre però la possibilità di effettuare alcune varianti del metodo MCMC, quali ad esempio il *reverse jump* o l'*exchange* MCMC usati da Courville et al. (2005). La mancanza di tale possibilità ha influito sulla decisione di usare una struttura fissa, in quanto il metodo MCMC offerto da WinBUGS si è rivelato non adatto ad esplorare l'unione di spazi di parametri con differenti dimensionalità che sarebbe risultata dal non avere una struttura fissa. La scelta di avere cinque cause fisse con archi verso tutti gli osservabili, sebbene più 'semplice' computazionalmente implica però l'abbandonare alcune caratteristiche del modello di Courville et al. (2005), tra cui parte della sua capacità di individuare 'modelli semplici' della realtà, ossia di spiegare gli eventi con poche cause ognuna legata solo ad alcuni osservabili. Per compensare

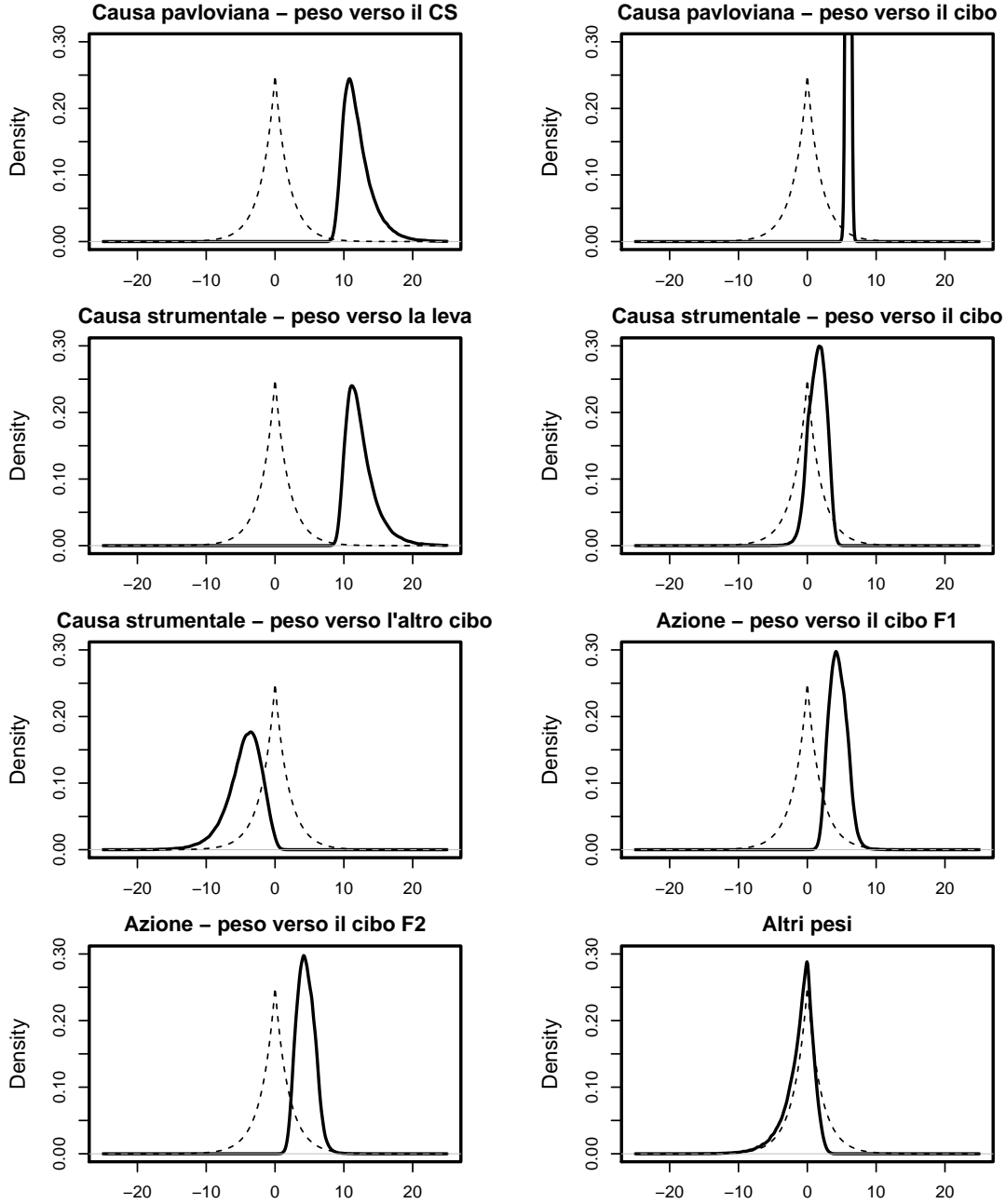


Figura 5.2: Distribuzioni *a posteriori* dei pesi. La curva tratteggiata mostra la distribuzione *a priori* (distribuzione di Laplace). Le cause pavloviane mostrano dei pesi positivi sia per il CS che per il cibo corrispondenti. Le cause strumentali mostrano invece pesi positivi verso la rispettiva leva e verso il rispettivo cibo. Il peso verso il cibo è solo leggermente positivo in quanto è necessaria la contemporanea influenza dell'azione per renderne possibile la presenza. Le cause strumentali sviluppano anche dei pesi negativi verso il cibo offerto dall'altra causa strumentale. L'azione di premere la leva ha influenza positiva su entrambi i cibi 'strumentali' F1 ed F2. Le distribuzioni degli altri pesi del modello rimangono concentrate su valori nulli.

ciò, abbiamo esplicitamente fatto considerare durante l'apprendimento la presenza di una sola causa per ogni vettore di osservazione. Come in Courville et al. (2005) la distribuzione a priori dei valori dei pesi era costituita da una distribuzione di Laplace (figura 5.2):

$$p(w_{ij}) = \frac{1}{2b} e^{-\frac{|w_{ij}|}{b}}, b = 2 \quad (5.2)$$

I pesi di *bias* degli osservabili sono stati invece fissati al valore di -6, ossia i vari osservabili tendono a non essere osservati nell'ambiente se non vi sono cause attive. Poiché nella distribuzione di Laplace usata la maggior parte della massa è concentrata nei dintorni dello zero, essa rappresenta quindi una credenza iniziale (*a priori*) secondo cui probabilmente le cinque cause nascoste non hanno nessun (o poco) effetto sugli osservabili. Le distribuzioni *a posteriori* dei pesi in figura 5.2 mostrano invece come a seguito dell'apprendimento il modello ha appreso che le cause nascoste hanno effetto ognuna su alcuni osservabili.

In particolare, le cause nascoste attribuibili all'apprendimento pavloviano (H1, H2, H3)² hanno una distribuzione *a posteriori* spostata verso pesi positivi nei confronti dei rispettivi suoni e cibi associati pavlovianamente. Se consideriamo quindi i valori più probabili dei pesi, le cause pavloviane hanno assunto una 'struttura' come quella in figura 5.3 (ignorando gli archi con pesi nulli).

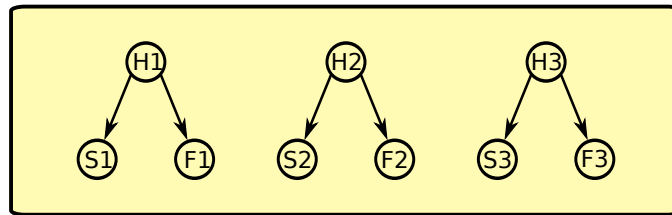


Figura 5.3: Fase Pavloviana: vengono appresi pesi positivi tra le cause latenti ed i suoni ed i cibi la cui presenza è correlata.

Per quanto riguarda le cause latenti strumentali invece la distribuzione a posteriori indica che è stata appresa un'influenza positiva delle cause sia sulla presenza delle leve che sulla presenza del cibo. Verso il cibo, questa influenza positiva è minore in quanto va associata all'influenza positiva dell'azione per far sì che effettivamente la presenza

²Poiché tutte le cause sono inizialmente connesse allo stesso modo con tutti gli osservabili, non è detto che al termine dell'apprendimento siano necessariamente le cause H1, H2 ed H3 ad essere associate agli stimoli pavloviani (suoni) e le cause H4 ed H5 agli stimoli strumentali (leve). È altrettanto possibile che siano le cause H2, H3, H5 ad essere scelte come 'cause pavloviane' e le restanti come strumentali. In ogni caso, il risultato è identico, al netto di un semplice 'cambio di nome' delle variabili (fenomeno del *label switching*).

del cibo possa essere osservata. È stata inoltre appresa una relazione inibitoria (peso negativo) tra ogni causa strumentale ed il cibo collegato all'altra causa strumentale. Questo peso negativo indica che la presenza di una causa strumentale ha un impatto negativo sulla possibilità di ottenere il cibo dell'altra causa strumentale. Detto in altre parole, se è presente la leva L1 del cibo F1, *non* è possibile ottenere il cibo F2 della leva L2. La 'struttura' appresa delle cause strumentali è quindi quella rappresentata in figura 5.4.

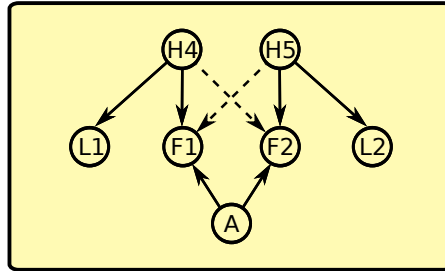


Figura 5.4: Fase strumentale: le cause strumentali (H4, H5), che danno luogo alla presenza delle leve (L1, L2), interagiscono con l'azione di premere la leva (A) nel rendere il cibo disponibile (F1, F2). Inoltre, ogni causa strumentale inibisce la disponibilità del cibo dell'altra leva (le linee tratteggiate indicano pesi negativi).

5.3 Effetto PIT

Una volta effettuata questa fase di apprendimento basata sull'esperienza pavloviana e strumentale, il modello è capace di spiegare i vari effetti PIT.

Durante la fase di test del PIT vi sono quattro possibili condizioni, che si differenziano a seconda della presenza o meno di differenti suoni assieme con una delle leve. Queste quattro condizioni, sulla base del modello appreso, danno origini ad inferenze diverse (vedi figura 5.5). In base a queste inferenze, cambierà anche la valutazione del ratto (o altro agente) riguardo alle conseguenze della sua possibile azione di premere la leva. Queste diverse valutazioni sono riconducibili ai vari effetti PIT descritti nei capitoli precedenti.

- **Baseline:** nella condizione di *baseline*, la leva L1 è presentata da sola, senza suoni. Sulla base della leva L1, il ratto può quindi inferire che la causa H4 è attiva. La presenza di H4, assieme all'azione A, può aumentare l'attivazione del cibo F1 e farlo diventare probabile. Il ratto, sapendo quindi di potere ottenere F1 tramite l'azione A, premerà quindi la leva con una certa frequenza.

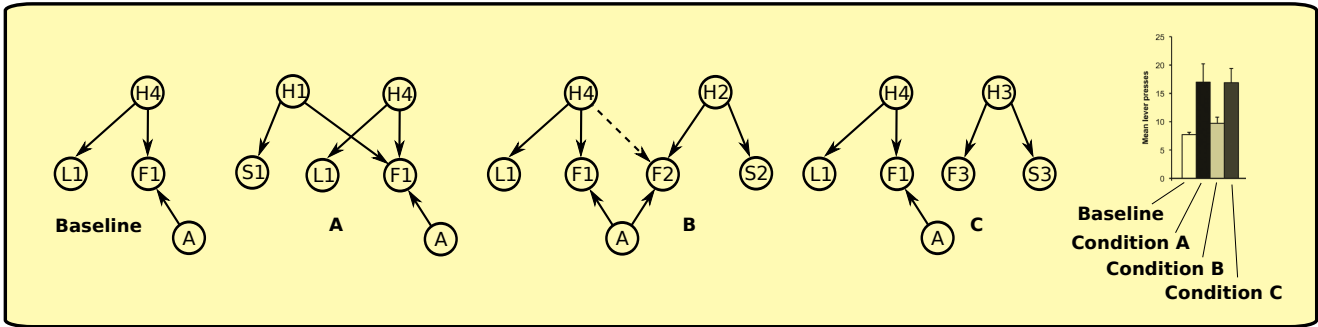


Figura 5.5: Fase di test del PIT: a seconda dei suoni presenti (S1, S2, S3 o nessuno), interazioni diverse si verificano nel modello. Solo i nodi rilevanti vengono mostrati in figura. Il modello può generare i tipici risultati sperimentali, come nell'istogramma (da Corbit & Balleine, 2011).

- Condizione A - PIT specifico:** dalla presenza del suono S1 e della leva L1, il ratto può inferire la presenza delle cause H1 ed H4. Entrambe le cause hanno un'influenza positiva su F1, ossia ne predicono la sua presenza nell'immediato futuro. Poiché ora l'attivazione di F1 è maggiore rispetto a quando vi è solo la leva L1, il ratto è più motivato a premere la leva in quanto vi sono *maggiori probabilità* di ottenere il cibo.
- Condizione C - PIT generale:** La presenza della leva L1 e del suono S3 implica che le cause H4 ed H3 sono presenti e che quindi i cibi F1 ed F3 sono possibili. Sebbene il cibo F3 non sia un effetto diretto dell'azione di premere la leva, la sua possibile presenza può in ogni caso motivare il ratto a premere la leva più che nella condizione *baseline*. Questa è chiaramente un tipo di motivazione diversa dalla Condizione A: anziché aumentare la probabilità di ottenere il cibo cui l'azione è diretta, viene 'aggiunta' una nuova ricompensa nello scenario dei risultati futuri a seguito dell'azione.
- Condizione B - assenza del PIT generale:** questa condizione è simile alla condizione C di PIT generale, ma in questo caso il caso cibo F2, evocato dal suono S2 attraverso la causa H2, è inibito dalla causa strumentale H4. A seguito della contemporanea attivazione (da parte di H2) ed inibizione (a causa di H4), il cibo F2 tende a rimanere attorno al suo valore base, ossia 'disattivo'. Solo il cibo F1 risulta quindi predetto come 'attivo' ed a conseguenza di ciò lo scenario diventa equivalente alla situazione di *baseline*. Non si osserveranno quindi differenze nel numero di pressioni della leva rispetto alla *baseline*.

Il modello esemplifica quindi come potrebbe funzionare l'ipotesi proposta nel capitolo 3. Il PIT specifico deriva da un'interazione tra cause latenti pavloviane e strumentali che risulta nella valutazione di una maggiore probabilità di ottenere la ricompensa collegata all'azione (*principio di efficacia*). Il PIT generale deriva invece dall'aggiunta di valore (una nuova ricompensa) nello scenario futuro a valle dell'azione, agendo quindi sul *principio di utilità*. L'assenza di PIT generale potrebbe invece essere conseguenza del fatto che la ricompensa 'predetta' dallo stimolo condizionato (S3) viene esclusa in quanto non possibile in presenza di H4 (*principio del contesto*).

6 Conclusioni

In questa tesi, abbiamo proposto una spiegazione funzionale del *Pavlovian-instrumental transfer* (PIT). Abbiamo avanzato un'ipotesi secondo la quale i due tipi di PIT noti in letteratura, PIT specifico e PIT generale, nonché l'assenza di PIT generale, sono associabili a 'tre principi dell'azione *goal-directed*'. La nostra idea è che questi tre effetti rappresentino l'azione degli stimoli condizionati pavloviani rispetto a tre aspetti della valutazione dell'azione che avviene nel caso di comportamento diretto ad un fine (*goal-directed*). Questi tre aspetti sono: il *contesto* (un'azione diretta ad un fine ha bisogno del giusto contesto), l'*efficacia* (un'azione può avere più o meno probabilità di conseguire un certo obiettivo) e l'*utilità* (le conseguenze dell'azione possono essere più o meno utili). Secondo questa ottica, il PIT specifico rappresenterebbe quindi l'utilizzo da parte dell'agente di indizi che indicano una maggior probabilità di ottenere una certa ricompensa. Dati a favore del coinvolgimento del PIT specifico con la probabilità di ottenere una ricompensa sono riscontrabili in Trick et al. (2011) in cui CS associati a maggiori probabilità di ricompensa elicitarono effetti PIT maggiori. Spiegare il PIT specifico in termini di probabilità, spiegherebbe inoltre perché il PIT specifico non risente della svalutazione della ricompensa (Rescorla, 1994; Holland, 2004; Corbit et al., 2007): la probabilità di ottenere una ricompensa è infatti indipendente dal suo valore. Al contrario, il PIT generale avrebbe influenza sull'utilità dell'azione, ossia sul suo valore, ed è infatti soggetto a svalutazione (Corbit et al., 2007). Il PIT generale corrisponderebbe all'utilizzo da parte dell'agente di indizi che indicano la presenza di una ricompensa aggiuntiva nell'immediato futuro e che quindi motivano l'agente ad eseguire le proprie azioni con più vigore. Infine, l'assenza di PIT generale corrisponderebbe all'utilizzo da parte dell'agente di indizi per comprendere quali ricompense siano disponibili in un dato contesto.

Abbiamo quindi creato un modello di come questa ipotesi funzionale possa essere tradotta in termini di computazione probabilistica. Nello specifico, abbiamo creato il nostro modello a partire da un lavoro di Courville et al. (2005) in cui il condizionamento pavloviano era spiegato nei termini di un modello generativo con cause latenti. Tramite

l'aggiunta di un nodo azione, abbiamo simulato il condizionamento strumentale; spiegando così i tre effetti PIT in termini di interazione tra le cause latenti alla base dei contesti pavloviani e strumentali creati dalle varie fasi di apprendimento del paradigma PIT.

Sebbene il modello al momento abbia una struttura di cause latenti fissa e non possa quindi essere considerato un'estensione diretta del modello di Courville et al. (2005) è in linea di principio possibile creare un modello che offra sia gli effetti PIT, sia gli effetti puramente pavloviani esposti in Courville et al. (2005).

In ogni caso, il modello attuale esemplifica la nostra ipotesi funzionale sui vari aspetti del PIT: da questa ipotesi è possibile trarre nuove predizioni tramite cui verificare il modello ed espandere la nostra conoscenza sul PIT. In particolare, poiché il PIT specifico risulta come un effetto di 'aumento delle probabilità di riuscita di un'azione', azioni strumentali che hanno già il 100% di probabilità di successo non dovrebbero essere soggette a PIT specifico (ma potrebbero essere soggette a PIT generale).

Abbiamo inoltre suggerito che il PIT specifico si possa verificare solo nell'ambito del comportamento *goal-directed*, cui la nostra ipotesi si riferisce, e che differenze riscontrabili nei dati in letteratura possano essere dovute all'utilizzo di programmi di rinforzo ad intervallo rispetto ai programmi *random-ratio*. Questo aspetto rimane da esplorare e da confermare.

Ci proponiamo in futuro di sviluppare un modello del PIT che non sia solo funzionale ma che sfrutti i numerosi dati sui substrati neurali del PIT per indagare anche i meccanismi tramite cui il PIT si verifica.

Infine, un'altra area da esplorare sono le interazioni tra PIT e droghe d'abuso: in particolare, l'effetto di *generalizzazione* delle droghe d'abuso (come discusso nel paragrafo 2.4.4) potrebbe essere spiegato da un modello che includa il funzionamento del PIT sia in ambito di comportamento *goal-directed* che di *habits*.

Bibliografia

- Balleine, B. W. & Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends in neurosciences*, 29(5), 272–9.
- Balleine, B. W. & Ostlund, S. B. (2007). Still at the choice-point: action selection and initiation in instrumental conditioning. *Annals of the New York Academy of Sciences*, 1104, 147–71.
- Beaumont, M. a. & Rannala, B. (2004). The Bayesian revolution in genetics. *Nature reviews. Genetics*, 5(4), 251–61.
- Blundell, P., Hall, G., & Killcross, S. (2001). Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *The Journal of Neuroscience*, 21(22), 9018–26.
- Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: conceptual foundations. *Trends in cognitive sciences*, 10(7), 287–91.
- Corbit, L. H. & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *The Journal of Neuroscience*, 25(4), 962–70.
- Corbit, L. H. & Balleine, B. W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *The Journal of Neuroscience*, 31(33), 11786–94.
- Corbit, L. H. & Janak, P. H. (2007a). Ethanol-associated cues produce general pavlovian-instrumental transfer. *Alcoholism, clinical and experimental research*, 31(5), 766–74.
- Corbit, L. H. & Janak, P. H. (2007b). Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *The Journal of Neuroscience*, 27(51), 13977–81.

Bibliografia

- Corbit, L. H. & Janak, P. H. (2010). Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *The European Journal of Neuroscience*, 31(7), 1312–21.
- Corbit, L. H., Janak, P. H., & Balleine, B. W. (2007). General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *The European journal of neuroscience*, 26(11), 3141–9.
- Courville, A. C., Daw, N. D., Gordon, G. J., & Touretzky, D. S. (2004). Model Uncertainty in Classical Conditioning. *Advances in Neural Information Processing Systems*, 16.
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2005). Similarity and discrimination in classical conditioning : A latent variable account. *Advances in Neural Information Processing Systems*, 17.
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in cognitive sciences*, 10(7), 294–300.
- El-Amamy, H. & Holland, P. C. (2007). Dissociable effects of disconnecting amygdala central nucleus from the ventral tegmental area or substantia nigra on learned orienting and incentive motivation. *The European journal of neuroscience*, 25(5), 1557–67.
- Gallistel, C. R. & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological review*, 107(2), 289–344.
- Gershman, S. J. & Niv, Y. (2012). Exploring a latent cause theory of classical conditioning. *Learning & behavior*, 40(3), 255–68.
- Hall, J., Parkinson, J. a., Connor, T. M., Dickinson, a., & Everitt, B. J. (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *The European journal of neuroscience*, 13(10), 1984–92.
- Hibbert, D. & Armstrong, N. (2009). An introduction to Bayesian methods for analyzing chemistry data. *Chemometrics and Intelligent Laboratory Systems*, 97(2), 211–220.
- Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of experimental psychology. Animal behavior processes*, 30(2), 104–17.

Bibliografia

- Holland, P. C. & Gallagher, M. (2003). Double dissociation of the effects of lesions of basolateral and central amygdala on conditioned stimulus-potentiated feeding and Pavlovian-instrumental transfer. *The European Journal of Neuroscience*, 17(8), 1680–1694.
- Holmes, N. M., Marchand, A. R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: a neurobehavioural perspective. *Neuroscience and biobehavioral reviews*, 34(8), 1277–95.
- Humphries, M. D. & Prescott, T. J. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Progress in neurobiology*, 90(4), 385–417.
- LeBlanc, K. H., Ostlund, S. B., & Maidment, N. T. (2012). Pavlovian-to-instrumental transfer in cocaine seeking rats. *Behavioral neuroscience*, 126(5), 681–9.
- Lex, A. & Hauber, W. (2008). Dopamine D1 and D2 receptors in the nucleus accumbens core and shell mediate Pavlovian-instrumental transfer. *Learning & memory*, 15(7), 483–91.
- Loredo, T. J. (1992). The promise of bayesian inference for astrophysics. In F. ED & B. GJ (Eds.), *Statistical Challenges in Modern Astronomy*, volume 297 (pp. 275–297). New York: Springer-Verlag.
- Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.
- Ostlund, S. B. & Balleine, B. W. (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(18), 4819–25.
- Ostlund, S. B., Maidment, N. T., & Balleine, B. W. (2010). Alcohol-Paired Contextual Cues Produce an Immediate and Selective Loss of Goal-directed Action in Rats. *Frontiers in integrative neuroscience*, 4(July), 1–8.
- Pearce, J. M. (1994). Similarity and discrimination: a selective review and a connectionist model. *Psychological review*, 101(4), 587–607.

Bibliografia

- Pezzulo, G., Rigoli, F., & Chersi, F. (2013). The mixed instrumental controller: using value of information to combine habitual choice and mental simulation. *Frontiers in psychology*, 4(March), 92.
- Pielock, S. M., Lex, B., & Hauber, W. (2011). The role of dopamine in the dorsomedial striatum in general and outcome-selective Pavlovian-instrumental transfer. *The European journal of neuroscience*, 33(4), 717–25.
- Poirier, D. J. (2006). The growth of Bayesian methods in statistics and economics since 1970. *Bayesian Analysis*, 1(4), 969–979.
- Prévost, C., Liljeholm, M., Tyszka, J. M., & O’Doherty, J. P. (2012). Neural correlates of specific and general Pavlovian-to-Instrumental Transfer within human amygdalar subregions: a high-resolution fMRI study. *The Journal of Neuroscience*, 32(24), 8383–90.
- Rescorla, R. a. (1994). Transfer of instrumental control mediated by a devalued outcome. *Animal Learning & Behavior*, 22(1), 27–33.
- Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In J. A. Black & W.F. Prokasky (Ed.), *Classical Conditioning II Current Research and Theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Robinson, T. E. & Berridge, K. C. (2003). Addiction. *Annual review of psychology*, 54, 25–53.
- Shiflett, M. W. (2012). The effects of amphetamine exposure on outcome-selective Pavlovian-instrumental transfer in rats. *Psychopharmacology*, 223(3), 361–70.
- Shiflett, M. W. & Balleine, B. W. (2010). At the limbic-motor interface: disconnection of basolateral amygdala from nucleus accumbens core and shell reveals dissociable components of incentive motivation. *The European journal of neuroscience*, 32(10), 1735–43.
- Solway, A. & Botvinick, M. M. (2012). Goal-directed decision making as probabilistic inference: a computational framework and potential neural correlates. *Psychological review*, 119(1), 120–54.

Bibliografia

- Trick, L., Hogarth, L., & Duka, T. (2011). Prediction and uncertainty in human Pavlovian to instrumental transfer. *Journal of experimental psychology. Learning, memory, and cognition*, 37(3), 757–65.
- Wyvell, C. L. & Berridge, K. C. (2000). Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward wanting without enhanced liking or response reinforcement. *The Journal of neuroscience*, 20(21), 8122–30.
- Wyvell, C. L. & Berridge, K. C. (2001). Incentive Sensitization by Previous Amphetamine Exposure : Increased Cue-Triggered Wanting for Sucrose Reward. *The Journal of neuroscience*, 21(19), 7831–7840.
- Yin, H. H. & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature reviews. Neuroscience*, 7(6), 464–76.