

Documents Tècnics de l'AFUNTAP

APUNTS BÀSICS DE CONDICIONAMENT OPERANT.

Mateu Servera Barceló

Departament de Psicologia. Universitat de les Illes Balears

<mateus@uib.es>

1. CONCEPTES BÀSICS.
2. PARADIGMES BÀSICS.
 - 2.1. Reforçament positiu o condicionament de recompensa
 - 2.2. Reforçament negatiu o condicionament d'omissió.
 - 2.3. Càstig positiu.
 - 2.4. Càstig negatiu o condicionament d'evitació/fugida.
3. FENÒMENS RELLEVANTS.
 - 3.1. Extinció
 - 3.2. Generalització i discriminació.
 - 3.3. Reforçament secundari
 - 3.4. Evitació i càstig.
3. PROGRAMES DE REFORÇAMENT.
4. CASOS ESPECIALS:
 1. Efecte de saciació.
 2. Programes de reforçament demorat i de reforçament diferencial de respostes.
 3. Principi de Premack.
 4. Efecte de contrast conductual

1. CONCEPTES BÀSICS.

L'origen del condicionament operant o instrumental són els treballs de Thordike i les seves lleis d'aprenentatge:

1. Llei del exercici: L'exercici o pràctica d'un tipus de resposta contribueix a enfortir la connexió entre ella i l'estímul que li correspon.
2. Llei de l'efecte (la més important): La resposta sempre està en funció de les conseqüències del medi (sense reforçament l'hàbit no es forma, no basta l'associacionisme).
3. Llei de la preparació o disposició: La preparació o disposició de l'organisme contribueix positiva o negativament a l'aprenentatge (reforçament intrínsec, expectatives, motivació...), és a dir, hi ha d'haver "*efecte novetat*".

Posteriorment fou SKINNER qui donà forma al COP impulsant el que seria després un model explicatiu per l'adquisició i manteniment de conductes (normals i anormals). El principi fonamental del COP és que la conducta està en funció de les conseqüències que la segueixen. Aleshores, al contrari del que passa en el condicionament clàssic o pavlovià (c. clàssic), els fets no són independents: la conseqüència és contingent a l'ocurrència de la resposta.

Terminologia bàsica del COP.

- CONDUCTA REFLEXE O RESPONENT (c. clàssic) = És un tipus de resposta autònoma (fisiològica) que normalment s'efectua sense aprenentatge previ, però que per un procediment de c. clàssic queda sota el control d'un determinant estímul antecedent que se'n diu "estímul condicionat" (EC). Davant aquest estímul sempre es produirà la conducta reflexe.

- CONDUCTA OPERANT = És un tipus de resposta normalment motora que té una taxa o probabilitat d'ocurrència controlada (al manco en gran part) per les seves conseqüències. Aquestes conseqüències se'n diuen "estímuls reforçadors" (ER).

- ESTIMUL REFORÇADOR (ER) = És qualsevol event físic present en el medi que aplicat contingentment a una conducta té la propietat d'alterar la seva ocurrència.

- REFORÇADOR POSITIU (ER+): És l'estímul (agradable) que aplicat contingentment a una conducta augmenta la seva probabilitat d'ocurrència:

- ER+ EXTERNES: primaris (estímuls no apresos, incondicionats..)
 secundaris (apresos, condicionats: materials, socials...)
- ER+ INTERNES (no observables: sentiment d'orgull, satisfacció..)

- REFORÇADOR NEGATIU (ER-): És un estímul (aversiu) que aplicat contingentment a una conducta disminueix la seva probabilitat d'ocurrència.

- ESTIMUL DISCRIMINATIU (ED) = És un estímul antecedent que assenyalava sempre la possibilitat d'aconseguir l'ER+ de manera contingent a una determinada conducta.

- ESTIMUL DELTA = És un estímul antecedent que assenyalava sempre la absència de l'ER+ de manera contingent a una determinada conducta.

- TAXA O FREQUÈNCIA DE RESPOSTA = És la forma més habitual d'avaluar la força d'una conducta operant. Consisteix en mesurar el nombre de vegades que l'organisme fa la conducta desitjada en un determinat interval de temps.

2. PARADIGMES BÀSICS.

Hi ha dos paradigmes o procediments bàsics en el COP:

1) EL REFORÇAMENT (REF) = És un procediment on sempre es provoca l'augment de la probabilitat d'ocurrència d'una determinada conducta.

Programa de reforçament positiu (REF+): aplicam un ER+ després de la conducta desitjada.

Programa de reforçament negatiu (REF-): retiram (o evitam) un ER- després de la conducta desitjada.

2) EL CÀSTIG = És un procediment on sempre es provoca una disminució de la probabilitat d'ocurrència d'una determinada conducta.

Programa de càstic positiu : aplicam un ER- després de la conducta no desitjada.

Programa de càstic negatiu: retiram (o eliminam) un ER+ després de la conducta no desitjada.

ESQUEMA DELS PARADIGMES DE CONDICIONAMENT OPERANT

	PRESENTACIÓ ER	SUPRESSIÓ ER	NO ER
Objectiu: Augmentar C.	REFORÇAMENT POSITIU o RECOMPENSA Presentar un ER+ després de la C. <i>Si fas..aleshores premi..</i>	REFORÇAMENT NEGATIU Suprimir o evitar un ER- després de la C. <i>Fes... i aleshores podràs deixar..</i> <i>Si fas... aleshores no hauràs de...</i>	
Objectiu: Disminuir C.	CÀSTIG POSITIU Presentar un ER- després de la C. <i>Si fas..aleshores càstig..</i>	CASTIG NEGATIU Suprimir un (o tots) ER+ després de la C. <i>Si fas...aleshores perds ..</i>	EXTINCIÓ (s'elimina l'ER+ que mantenia la C.)

Nota: ER Estímul Reforçador (+, positiu ó -, negatiu). C.=conducta

Exemples i aclariments:

1. REF POSITIU (REF+): Qualsevol procediment pel qual el subjecte rebí un ER+ per efectuar una resposta desitjada (que es vol crear o incrementar).

Exemple: cada vegada que acabis els exercicis marcats en el temps indicat pots jugar a X.

Nota: *És molt important que X sigui qualsevol joc novedós o molt especial per a l'infant, i que a més no pugui accedir-hi de cap altra manera que no sigui fent els exercicis. Un ER+ ha de ser novedós i escàs en la vida de la persona perquè incrementi el seu efecte.*

2. REF NEGATIU (REF-): La conducta operant té com a objectiu o bé eliminar la presència d'un ER- ó bé impedir la seva aparició. En el primer cas se'n diu **CONDICIONAMENT DE FUGIDA**, mentre en el segon cas **CONDICIONAMENT D'EVITACIO**. En ambdós casos l'objectiu és incrementar la freqüència de la conducta.

Exemples:

Fugida: No et pots aixecar d'aquesta taula fins que acabis els deures. No et deixaré jugar a res, ni fer res fins que desis la teva habitació.

Evitació: Si ajudes a parar taula no hauràs de rentar plats. Si fas els deures cada dia no hauràs de quedar-te a casa a estudiar els dissabtes.

3) **CÀSTIG POSITIU**: És el càstig tal i com s'entén col·loquialment. S'aplica un ER- de manera contingent a una conducta que es vol eliminar.

Exemple: Cada vegada que el nen pega al seu germà petit reb un crit, una renyada, un cop a les anques. etc.

Nota: *el principal inconvenient d'aquest procediment és que perquè sigui eficaç l'ER- ha de ser potent, d'elevada intensitat, i això crea molts de problemes ètics i secundaris (estats de ràbia i rancúnica en el nen, aprenentatge de conductes agressives per modelat, etc.). Si l'ER- és de baixa intensitat aviat el nen s'hi acostuma i perd eficàcia.*

4) **CÀSTIG NEGATIU**: És produeix quan, de forma contingent, l'emissió d'una conducta no desitjada suposa la pèrdua d'un ER+ que ja estava en possessió del nen (**COST DE RESPOSTA**), la pèrdua d'un ER+, o part d'ell, que es preveia que es podria obtenir (**OMISSIÓ**), ó la retirada total o parcial dels ER+ presents en el medi (**TEMPS FORA**).

Exemple:

Cost de resposta: Cada vegada que el nen diu una paraulota, un insult, l'infant ens ha de donar 20 cèntims de la seva guardiola.

Omissió: Cada vegada que el nen diu una paraulota, insult, són 20 cèntims menys a cobrar de la paga setmanal.

Temps fora (TF): Cada vegada que, quan jugam en situació de grup amb altres nens, el nostre nen s'enfada i insulta o agredeix als altres és confinat temporalment en una habitació (sense ER+) o bé ens anem tots i el deixam totsol (TF parcial).

3. ALTRES PROCEDIMENTS RELLEVANTS.

3.1. EXTINCIÓ.

És defineix com a la supressió de l'ER+ que durant un temps ha estat el responsable del manteniment de la conducta que es vol eliminar. A la llarga això produeix una disminució gradual de la resposta apresada i un augment de les respostes alternatives que teòricament poden competir amb ella. Però inicialment sol tenir un efecte advers: l'augment de la taxa de la conducta problema. Normalment això dura poc temps, i acaba cedint, però si hi ha històries prèvia d'aprenentatge on el nen ha estat sotmés a procediments similars que no s'han aplicat correctament pot produir-se una resistència important.

D'altra banda, és normal que de tant a tant, quan el procediment d'extinció ja està funcionant, torni a aparèixer puntualment la conducta problema, o fins i tot que es mantengui encara que sigui en una taxa molt més baixa que abans.

3.2. REFORÇAMENT PARCIAL

El principi bàsic és que els subjectes que són reforçats parcialment són més resistents a l'extinció que els subjectes que són reforçats de forma continua. És a dir, la força de la resposta no té perquè augmentar en funció del nombre de vegades que és reforçada.

3.3. CONTROL ESTIMULAR I REFORÇAMENT DIFERENCIAL: GENERALITZACIÓ I DISCRIMINACIÓ.

Es diu que hi ha CONTROL ESTIMULAR quan en presència d'un determinat ED augmenta molt la probabilitat de que aparegui un tipus de resposta en particular i no se'n facin d'altres. El control estimular és realment el que permet i afavoreix l'aprenentatge (permet respondre quan toca i no respondre quan no toca). Això va molt lligat a dos processos:

La GENERALITZACIÓ D'ESTÍMULS pot definir-se com el procés pel qual els organismes fan les mateixes conductes apresades davant ED específics en presència d'altres estímuls similars però diferents a l'original. És un principi fonamental de l'aprenentatge que fa que puguem respondre adequadament en moltes situacions on mai abans havíem rebut entrenament per fer-ho. També és possible la generalització de respostes: quan un estímulo pot provocar una determinada resposta, també és capaç de provocar respostes molts semblants o relacionades sense entrenament previ.

La DISCRIMINACIO (totalment unida a la generalització) consisteix en que una determinada resposta es veu reforçada davant un determinat estímul, però no davant un altre encara que hi guardi certa similitud.

Generalització i discriminació són part fonamental del REFORÇAMENT DIFERENCIAL i també permeten l'entrada en funcionament dels ER condicionats o secundaris. Sens dubte aquests procediments són els que més incidència tenen, pel que fa al COP, en l'aprenentatge del comportament humà. El repertori conductual de la persona es fa formant des de petits, a través dels ER+, ER- o extincions que reben les nostres conductes per part dels altres. En general parlem de ER de caire social que són el més freqüents i amb molts de casos el que major efecte tenen sobre la conducta dels infants.

3.4. REFORÇADOR SECUNDARI I REFORÇAMENT CONDICIONAT

Un reforçador secundari és un estímul originalment neutre que mitjançant el seu aparellament amb un reforçador primari adquireix les seves propietats, de forma que, en posteriors ocasions, per sí mateix reforçarà la conducta. Els diners i l'alabança social són dos dels reforçadors secundaris més poderosos.

3. PROGRAMES DE REFORÇAMENT.

Bàsicament podem destacar PROGRAMES DE REFORÇAMENT CONTINU (cada resposta és reforçada immediatament) i PROGRAMES DE REFORÇAMENT PARCIAL (que, com abans hem dit, garanteixen una força major de la resposta al ser més resistents a l'extinció).

Els programes de REFORÇAMENT PARCIAL s'apliquen sota dos parametres: TAXA DE RESPOSTA (nombre de respostes necessàries per obtenir el reforçador) I TEMPS (temps que passarà fins aconseguir el reforçador). A partir d'aquesta distinció distingim 4 tipus de programes simples:

1) RAÓ FIXA: el subjecte és recompensat cada N respostes correctes efectuades (la taxa és fixa). És a dir, després d'haver efectuat un nombre determinat de resposta, la pròxima es reforçarà (RF5, cada 5 respostes cau el reforçador). La taxa de resposta és alta i estable però amb pausa llarga. Hi ha poca resistència a l'extinció.

Exemple: Cada tres exercicis acabats guanyes 5 minuts de joc a la vídeo-consola.

2) RAÓ VARIABLE: el subjecte és recompensat després d'un nombre variable de respostes emeses. L'experimentador determina una mitjana de respostes (p.e, RV5, la mitjana és de cinc, però a vegades es reforça després de 1, de 7, de 3, o de 9 conductes). La taxa de resposta és molt alta i estable amb molt poques pauses. Augmenta la resistència a l'extinció.

Exemple: Mentre el nen va fent exercicis li donam, segons el programa estipulat (p.e. RV5), tiquets per valor de 5 minuts de vídeo-consola. A vegades el rebrà després d'un exercici i a vegades després de fer-ne set. La qüestió és tenir-ho programat durant l'hora de deures que ha de fer el nen.

3) INTERVAL FIXE: el subjecte reb el reforçador a la primera resposta efectuada després d'un interval fixe de temps (IF300", després de 300 segons la primera resposta correcta és reforçada). La taxa de resposta davalla després del reforçador i a mesura que passa el temps va augmentant progressivament. L'extinció és similar a la RF.

Exemple: Després de que el nen dugui 5 minuts fent feina, en acabar el primer exercici reb un tiquet vàlid per 5 minuts de joc a la vídeo-consola.

4) INTERVAL VARIABLE: el subjecte reb el reforçador a la primera resposta efectuada després d'un interval variable de temps (IV300", la mitja és 300 segons, però el pot rebre als 10 o als 900 segons). La taxa és estable però no molt alta. Molta resistència a l'extinció.

Exemple: Mentre el nen va fent exercicis li donam, segons el programa estipulat (p.e. IV300"), tiquets per valor de 5 minuts de vídeo-consola després del primer exercici acabat cada cert temps de feina. A vegades aquest temps de feina podran ser 30" i a vegades 10 minuts. La qüestió és tenir-ho programat durant l'hora de deures que ha de fer el nen.

Els programes complexos són combinacions de dos o més programes simples. Per exemple: en un programa complex de IF5' i RF3 passarà que el reforçador s'administrarà després de la tercera resposta observada dins un interval de temps de cinc minut.

Exemple: El nen rebrà el tiquet de joc a la vídeo-consola cada vegada que acabi 3 exercicis en 5 minuts.

La conducar de "zapping" és un dels millors exemples de programa complex de RV i IV: a vegades basta activar el TV i ja trobes un programa que et satisfà. A vegades no cal respondre més (canviar de canal) però has d'esperar un temps perquè acabi la publicitat (o el que sigui) i veure el teu programa (falta saber, però, si aquell dia serà o no del teu gust). En altres ocasions has de premer el botó dels canals un nombre variable de vegades, durant un temps variable... En definitiva, estam sotmesos a un programa complex, variable en raó i interval, i això provoca que normalment la conducta de zapping sigui de freqüència molt elevada, molt estable i molt resistent a l'extinció.

4. CASOS ESPECIALS:

1. Efecte de saciació.

És un fenomen molt relacionat amb l'extinció. De manera relativament sorprenent consisteix en presentar o mantenir el reforçador d'una conducta problema durant períodes de temps llargs. L'efecte que produeix és que el reforçador perd el seu valor inicial i la conducta s'acaba extingint. L'explicació d'aquest efecte pot tenir relació amb un procediment de c. clàssic: el reforçador ó EC (Estímul Condicionat) s'acaba convertint en un EC aversiu a causa del seu excés (esgota al subjecte), i acaba per provocar una RC (resposta condicionada) de defensa. Pot tenir utilitat en conductes addictives: fumar, menjar dolços, jugar a la vídeo-consola, etc.

2. Programes de reforçament demorat.

Un dels principis bàsics del COP és que els ER han de ser contingents. Això normalment ha de ser així al principi però amb el temps és fonamental que es vagi introduint una certa demora. En la vida real hem de fer moltes conductes durant llargs intervals de temps abans de rebre el reforçador (el cas de l'estudiant amb un examen a la fi del curs). La consistència temporal en la conducta la fan humans, però també la fan animals (alguns tarden setmanes en fer un niu o mesos en vols migratoris) i és, en definitiva, un dels objectius fonamentals de l'educació dels infants. Com més aviat aprenen els nens a ser constants i amb capacitat de demora dels reforçadors més aviat es pot dir que “maduren” i tenen molt més probabilitats d'èxit acadèmic i social, mentre en cas contrari tenen moltes més probabilitats de sentir frustració i tenir alteracions de comportament (qualsevol ser viu que no se li proporciona reforçador quan l'espera experimenta reaccions emocionals adverses). L'ensenyament d'això és pot fer a través de programes de reforçament demorat. La demora sempre té un caràcter progressiu combinant taxa de respostes i temps. Així, per exemple, inicialment un nen pot rebre reforçament només per intentar resoldre un exercici, amb el temps el reforçament passa a la finalització de l'exercici, després serà cada dos, tres, cinc i finalment pot ser a la finalització de tota la tasca.

Nota: el reforçador sempre ha d'aparèixer a la fi d'una seqüència perquè sinó podem caure en un procés d'extinció. La clau de la demora està, probablement, en que cada petita seqüència que una vegada va ser reforçada es converteix al mateix temps en ED i ER+ per la següent seqüència i així progressivament. D'Això, tècnicament, se'n diu MOLDEJAMENT si el que estam treballant són conductes noves per a l'infant, o ENCADENAMENT si el que estam fent és millorar conductes que l'infant ja sap fer però no massa bé.

4. Programes de reforçament diferencial.

El *reforçament diferencial d'altres conductes* (RDA o, a vegades en els textos en castellà, RDO) És un procediment mitjançant el qual en una determinada situació el reforçador pot seguir a un conjunt o a totes les respostes del subjecte, excepte en una que es pretén eliminar. És un principi fonamental que ve a dir que per eliminar una conducta, no només s'ha de mirar d'extingir-la, sino que també convé reforçar-ne d'altres d'alternatives. Estrictament parlant el RDA és basa en un castig negatiu d'omissió, on, com sabem, fer una conducta X elimina la possibilitat d'accedir a un ER+ i, per lògica, totes les altres conductes es veuran reforçades per sí mateixes. Ara bé, un RDA és una mica més refinat, perquè el procediment de reforçament positiu sol estar molt més delimitat (no és tan general), de manera que a més d'estar clar quina conducta inhibirà el ER+, també està clar quina ho permetrà.

Per exemple: Un mestre posa en marxa un programa de IF15' que consisteix en que cada 15 minuts una senyal l'avis i atorga 2 minuts de temps lliure a tots els alumnes que estan asseguts. No està en el teu lloc suposa la pèrdua d'aquest premi. Al cap del dia, pot haver-hi alumnes que hagin acumulat un màxim de 60 minuts i la darrera hora la tenen lliure per fer el que vulguin (dintre d'un cert ordre preestablert).

Per tant, podem dir que un RDO implica certament el castig negatiu per ommissió: no està assegut impedeix un ER+, *pero probablement no és el que controla la conducta*. El controlador està en el reforçament positiu que deriva del fet de seure (facis el que facis mentre seus, perquè de moment d'això no ha estipulat res el mestre... si bé amb el temps podria afegir-hi més condicions).

El reforçament diferencial de respostes de baixa taxa (RDB) és quan es reforça una conducta només si durant un interval de temps no supera un determinat nombre d'ocurrències. La baixa taxa pot ésser definida a partir de la "taxa promig" en un interval de temps determinat, o bé a partir de "taxa d'operant a operant" (més recomanable), és a dir, a partir de l'interval de temps que ha de passar necessàriament entre una resposta i l'altra perquè sigui reforçada. En aquest cas, a l'igual que abans, estam combinant reforçament positiu amb càstig negatiu d'omissió (sobrepasar la freqüència de la conducta impedirà tenir el reforçador, però mantenir-se dins la taxa permet l'accés). No es pot parlar de presència de l'extinció en sentit tècnic perquè no hem identificat clarament el reforçador que manté la conducta per després suprimir-lo i, d'altra banda, hem de recordar que l'extinció tendria com a objectiu suprimir del tot la conducta, però no és el cas.

Hi ha moltes conductes que haurien de ser de baixa taxa, és a dir, s'haurien de decrementar però no suprimir (el nen que *s'enrolla* massa en un examen, que parla massa a classe, que puja massa vegades a la xarxa jugant a tennis, etc.). L'objectiu és aconseguir que passi un temps entre resposta i reposta.

Exemples: El mestre reforça la participació en els debats de classe, però només si és produïxen entre una i tres intervencions cada 10 minuts (s'ha d'intervenir però s'ha de deixar parlar als altres). El mestre para atenció i reforça als alumnes que durant la realització d'un exercici li demanen que revisi la seva feina entre una i tres vegades (es considera important que el nen consulti coses però no tota l'estona, que aprengui a seleccionar allò que realment sembla més important)

3. El principi de Premack.

El principi que estableix que l'accés contingent a conductes d'alta freqüència serveix com a reforçador per a efectuar conductes de baixa freqüència. Col·loquialment es conegut com la "*lleï de l'àvia o de la padrina*" i el que fa es escollir activitats que són molt freqüents i agradables per a l'infant i posar-les en relació amb d'altres activitats que ha de fer i que li costen més. Per exemple, pots jugar un temps amb la vídeo-consola després de sopar si ajudes a desparar la taula, et fas les dents netes i et poses el pijama

4. L'efecte de contrast conductual .

El contrast conductual consisteix en que si una conducta es manté en dues o més situacions i introduïm un procediment per fer-la disminuir en una d'aquestes situacions, l'efecte serà un augment de la seva freqüència en les altres situacions. Per exemple: prohibir fumar en casa pot provocar un augment d'aquesta conducta fora de casa.