

Apprentissage du comportement optimal par exposition aux contingences dans le dilemme de Monty Hall

(Learning optimal behavior through exposure to contingencies in the Monty Hall dilemma)

Matthieu Villatte, David Scholiers et Esteve Freixa i Baqué¹

Université de Picardie Jules Verne.Département de Psychologie.Laboratoire
E.C.C.H.A.T.

Depuis l'antiquité, le thème de la rationalité a donné lieu à de nombreux débats et recherches en philosophie et en psychologie. Deux approches peuvent être distinguées : l'approche descriptive, qui s'attache à l'étude empirique de la façon dont raisonne la plupart d'entre nous dans la vie quotidienne, et l'approche normative, dont le but est de découvrir les règles canoniques d'un raisonnement correct (Stich, 1997, 1998, 1999) . A partir de ces études, les recherches ont rapidement conjugué ces deux approches afin d'examiner dans quelle mesure les individus seraient aptes ou non à produire des jugements rationnels. Si, pour Aristote, l'Homme était un animal rationnel par essence, et si, plus récemment, pour Goodman (1984 - cité par Stich, 1999 -) et Cohen (1981 - cité par Shier, 2000 -), on ne peut démontrer l'irrationalité de l'Homme, car les critères de rationalité se basent sur un catalogue des intuitions de la population générale, des études de psychologie expérimentale ont pourtant mis en évidence depuis quelques décennies des erreurs systématiques de jugement en situation d'incertitude, telles que le problème des quatre cartes de Wason et Johnson-Laird (1970), la négligence du taux de base de Kahneman, Slovic et Tversky (1973) ou l'erreur de conjonction de Kahneman et Tversky (1982). Dans ce type de problème, une très large majorité des individus émet des choix qui ne sont pas conformes aux règles de probabilités.

¹Université de Picardie Jules Verne. Département de Psychologie. Laboratoire E.C.C.H.A.T.- Chemin du Thil. 80025 AMIENS Cedex 1. adresse électronique: matthieu.villatte@u-picardie.fr, estev.freixa@u-picardie.fr

Reprenant la distinction formulée initialement par Chomsky (1967) entre compétence et performance, les sciences cognitives ont développé des théories explicatives de ces erreurs systématiques de jugement en termes de défaut des principes de raisonnement internalisés. Un comportement de choix erroné serait alors le produit de l'emploi de règles internes (ou heuristiques) efficaces dans certains domaines, mais pas pour la situation particulière dans laquelle est apparu ce comportement.

Certains chercheurs, tel Palmarini (1993), ont développé le concept d'illusion cognitive, qui désigne un phénomène générant un comportement irrationnel et lui-même dû à l'action d'un « inconscient cognitif » biaisant le jugement. Les processus mentaux qui provoqueraient le comportement inadéquat fonctionneraient de façon similaire aux processus de perception visuelle incontrôlables par l'individu dans les cas d'illusions d'optique : c'est-à-dire que l'erreur de jugement serait irrépessible en situation d'illusion cognitive.

En revanche, la recherche en Analyse Expérimentale du Comportement, qui a pour fondement le postulat selon lequel les comportements sont, pour la plupart, produits par apprentissage, a montré qu'une erreur systématique de jugement telle que la négligence du taux de base pouvait être appréhendée comme un comportement non-optimal acquis: même si le choix inadéquat émis par les sujets dans le cadre de ce problème semble très résistant à un apprentissage du choix optimal par essais renforcés (Goodie et Fantino, 1996 - voir aussi Stolarz-Fantino, Fantino, Zizzo et Wen, 2003 -), on peut avancer l'hypothèse selon laquelle cette erreur est un phénomène issu d'associations antérieures acquises et grandement dépendant du contexte dans lequel il survient, puisqu'en prolongeant les séries d'essais, le choix non-optimal peut diminuer significativement (Fantino 1998).

Scholiers, Villatte et Freixa i Baqué (2003) ont également montré que l'erreur de jugement probabiliste survenant systématiquement dans le dilemme de Monty Hall pouvait diminuer significativement en modifiant les contingences de renforcement. Dans ce dilemme, un participant doit choisir entre trois gobelets retournés afin de découvrir une pièce cachée sous l'un d'eux. Une fois ce choix effectué, et avant de soulever le gobelet choisi afin de vérifier si le sujet a gagné ou perdu, un gobelet vide et non choisi est retiré. Le participant a alors la possibilité soit de conserver son choix initial, soit de le modifier pour choisir l'autre gobelet restant. Le taux de choix pour la stratégie « Modifier » est en général en dessous de 20 % à l'exposé du problème (Granberg et Brown, 1995 ; Aaron et Spivey, 1998 ; Friedman, 1998 ; Granberg et Dorr, 1998 ; Mueser et Granberg, 1999 ; Slembeck et Tyran, 2002 ; Bown, Read et Summers, 2002 ; Franco-Watkins, Derks et Dougherty, 2003 ; Tubau et Alonso, 2003). Un taux similaire a été observé dans des cultures différentes - au Brésil, en Suède, en Chine et

aux Etat Unis - par Granberg (1999). Par ailleurs, d'après Krauss et Wang (2003), toutes les études expérimentales concernant le dilemme de Monty Hall montrent, jusqu'à maintenant, les mêmes résultats : la grande majorité des participants pense, à l'énoncé du problème, que « Modifier » et « Conserver » ont des probabilités de gain équivalentes. Pourtant, les règles de probabilités nous indiquent que la stratégie « Modifier » a une probabilité de gain deux fois supérieure à celle de la stratégie « Conserver ».

Scholiers, Villatte et Freixa i Baqué (2003) ont fait varier le nombre de gobelets constituant le dilemme de Monty Hall (3, 5 et 10 gobelets), modifiant ainsi les taux de renforcement associés à chaque stratégie (1/3 versus 2/3 ; 1/5 versus 4/5 ; 1/10 versus 9/10), et ont soumis les participants à l'observation répétée des conséquences de chaque stratégie. Ils ont obtenu une augmentation significative de la probabilité d'apparition du comportement optimal, c'est-à-dire, du choix de la stratégie « Modifier », en fonction de l'augmentation du taux de renforcement de la stratégie optimale.

Bien que l'hypothèse selon laquelle l'erreur de jugement observée dans le dilemme de Monty Hall est un comportement non-optimal produit par apprentissage semble confirmée par ces résultats, plusieurs critiques peuvent être formulées à l'égard de cette étude.

Premièrement, même si la modification des contingences environnementales, dans le but de modifier un comportement, est une démarche cohérente avec le postulat béhavioriste selon lequel les comportements sont produits et maintenus par les contingences, on peut se demander si, en faisant varier le nombre de gobelets, les caractéristiques du dilemme de Monty Hall n'ont pas été changées au point de transformer la nature même du problème.

Deuxièmement, le fait d'avoir soumis les sujets à une observation des contingences n'a pas été efficace pour diminuer significativement la probabilité d'apparition du comportement non-optimal dans le dilemme de Monty Hall présenté avec trois gobelets. On peut alors se demander si une exposition concrète aux conséquences du choix serait plus efficace pour l'apprentissage du comportement optimal. Cette hypothèse a été examinée par Friedman (1998), par Slembeck et Tyran (2002), et par Palacios-Huerta (2003). Ces auteurs ont soumis les participants au dilemme de Monty Hall durant des séries de 10 à 40 essais renforcés et ont constaté une diminution significative, mais très faible, de la probabilité d'apparition du comportement non-optimal, et les taux observés sont très éloignés des 100 % qui correspondent au comportement optimal.

Comment, alors, les contingences pourraient-elles être aménagées de façon à augmenter la probabilité d'apparition du comportement optimal sans pour autant modifier les caractéristiques du dilemme de Monty Hall ?

Premièrement, on peut émettre l'hypothèse selon laquelle ajouter une valeur aversive

aux cas de perte (c'est-à-dire, lorsque la stratégie choisie par le participant ne permet pas de trouver la pièce) permettrait d'augmenter la prégnance des contingences. En effet, dans l'étude de Scholiers, Villatte et Freixa i Baqué (2003), l'absence de renforcement en cas de perte (deux fois sur trois) a pu être moins efficace pour éteindre le comportement non-optimal que le renforcement (une fois sur trois) pour le maintenir, selon un mécanisme similaire à celui produisant un comportement superstitieux (Catania et Cutts, 1963).

Par ailleurs, certaines études expérimentales montrent que lorsque le dilemme de Monty Hall est présenté sous la forme d'un jeu avec deux adversaires (un premier sujet choisit une carte parmi trois retournées afin de trouver où se trouve l'as, les deux autres cartes revenant à un deuxième sujet qui sait où se trouve l'as ; le deuxième sujet retourne alors une de ses cartes n'étant pas l'as et le premier sujet doit choisir de conserver son choix initial ou de le modifier et de choisir la carte restante du deuxième sujet), l'efficacité de l'apprentissage par essais répétés s'en trouve renforcée, c'est-à-dire que le nombre de stratégies « Modifier » choisies augmente (Tubau et Alonso, 2003). Il en va de même lorsque les sujets expérimentaux sont en compétition, c'est-à-dire, chaque sujet participant au jeu seul face à un ordinateur et devant effectuer un score supérieur aux autres (Slembeck et Tyran, 2002).

Dans ces situations expérimentales, il est probable que la valeur aversive de l'échec soit supérieure à celle observée dans l'expérience de Scholiers, Villatte et Freixa i Baqué (2003) ; en effet, à chaque fois que le sujet ne gagne pas avec la stratégie qu'il a choisie, c'est un autre qui gagne dans la présentation du dilemme sous forme d'un jeu avec des adversaires ; et dans la situation de compétition, à chaque échec, le sujet prend du retard dans la concurrence avec les autres participants.

Cependant, les travaux de Dinsmoor (1983 - cité par Catania, 1994 -) ont montré que les individus produisent des stimulus discriminatifs lorsque ceux-ci sont corrélés avec des renforçateurs appétitifs, mais qu'il n'en va pas de même lorsqu'il s'agit d'événements aversifs. Ainsi, l'efficacité d'un message dépendrait d'avantage de son caractère appétitif ou aversif pour le récepteur que de son contenu informatif. L'association entre un stimulus appétitif et un stimulus aversif apportant la même information peut-elle, alors, s'avérer plus efficace pour l'apprentissage du comportement optimal qu'un seul de ces stimulus isolé ?

Deuxièmement, pour que les stratégies « Conserver » et « Modifier » aient des taux de renforcement qui soient significativement dans les limites attendues, il apparaît utile d'augmenter le nombre d'essais employés jusqu'ici. L'efficacité de l'exposition concrètes aux contingences pourrait se trouver augmentée par une extension à 100 essais.

MÉTHODE

L'expérience consiste en une exposition prolongée aux contingences concrètes dans le dilemme de Monty Hall, c'est-à-dire, en une série de 100 parties à choix de stratégie libre. Cette expérience est constituée de trois conditions : une condition dans laquelle la procédure d'apprentissage est basée sur le renforcement positif : RP (à chaque fois que la stratégie choisie – « Conserver » ou « Modifier » – permet de gagner, on présente un stimulus appétitif) ; une condition dans laquelle la procédure d'apprentissage est basée sur la punition négative : PN (à chaque fois que la stratégie choisie – Conserver ou Modifier – ne permet pas de gagner, on retire un stimulus appétitif) ; et une condition dans laquelle la procédure d'apprentissage est basée sur le renforcement positif et la punition négative : RP + PN.

Les sujets effectuent les parties à l'aide d'un logiciel informatique qui reproduit exactement les conditions du dilemme de Monty Hall présenté avec des gobelets concrets, sauf que la pièce est remplacée par une balle. Cependant, l'aspect virtuel de la présentation du jeu (contrairement à l'aspect concret des gobelets réels) peut induire chez les sujets le sentiment que ce jeu est truqué, d'autant plus que les conséquences de ses choix vont avoir tendance à contredire son opinion première concernant la meilleure stratégie à employer. Afin de limiter au maximum les doutes éventuels à ce sujet, un règlement du jeu officiel (certifié par le Régisseur d'avances de la Faculté de Philosophie, Sciences Humaines et Sociales de l'Université de Picardie Jules Verne) est présenté au sujet, stipulant que le logiciel reproduit exactement les conditions du jeu concret et que l'emplacement de la balle est tiré au sort par l'ordinateur pour l'ensemble des 100 parties. Cet emplacement figure sur une feuille qui est fournie au sujet avant de jouer mais qu'il ne peut consulter qu'à la fin des 100 parties, pour vérifier la correspondance entre le tirage au sort préalable et l'emplacement de la balle qui a été effectivement observé.

Toutefois, pour l'énoncé du problème avant la procédure d'apprentissage, on utilise des gobelets réels.

Participants

45 participants âgés de 18 à 28 ans tirés au hasard et répartis en 3 groupes indépendants de 15 participants. Les participants, étudiants volontaires de 1^{ère} année de psychologie, ont été rétribués d'un demi point sur leur moyenne du module de méthodologie expérimentale de deuxième année.

Matériel

- 3 gobelets opaques et identiques (pour la présentation du jeu).
- 1 jeton blanc pour marquer le gobelet choisi par le sujet (pour la présentation du jeu).
- Une feuille d'engagement à faire signer à chaque sujet, stipulant que celui-ci accepte de ne révéler aucune information concernant l'expérience à tout personne susceptible d'y participer ultérieurement. Cet engagement n'a aucune valeur administrative ou juridique, mais, conformément à ce que nous apprennent les travaux de Joule et Beauvois (1987, 1998) sur la théorie de l'engagement, il permet d'augmenter les chances que le déroulement de l'expérience reste secret, ce qui est particulièrement important dans ce cas précis, puisqu'il s'agit d'un problème à résoudre.
- Le règlement du jeu officiel certifié par le Régisseur d'avances de la Faculté de Philosophie, Sciences Humaines et Sociales de l'Université de Picardie Jules Verne.
- Un ordinateur connecté à internet.
- Le logiciel informatique hébergé sur un site internet contenant les fonctionnalités suivantes :

Une partie publique avec à l'écran trois gobelets alignés à l'envers ; un bouton blanc sous chaque gobelet ; deux boutons « Conserver » et « Modifier » ; un bouton « OK » de validation de choix ; le score ; le nom et le prénom du sujet.

Au début de chaque partie, apparaît d'abord l'instruction suivante : « Choisissez un gobelet et validez ». Le sujet choisit un gobelet en cliquant sur l'un des trois boutons blancs se situant sous les gobelets, celui devient alors rouge. Il valide ensuite son choix en cliquant sur le bouton « OK ». Un gobelet qui n'a pas été choisi et qui ne contient pas la balle disparaît de l'écran et une nouvelle instruction apparaît : « Un gobelet ne contenant pas la balle a été retiré. Choisissez de conserver ou de modifier votre choix initial et validez ». Le sujet clique sur le bouton « Conserver » ou sur le bouton « Modifier » puis clique sur le bouton « OK ». Le gobelet contenant la balle est soulevé et la mention « Perdu ! » ou « Gagné ! » apparaît selon que la stratégie choisie a permis ou non de trouver la balle. L'instruction « Nouvel essai ? » apparaît et le sujet clique sur le bouton « OK ». Le score évolue (un point supplémentaire si la balle a été trouvée avec la stratégie choisie dans les conditions renforcement positif et renforcement positif plus punition négative ; un point en moins si la balle n'a pas été trouvée avec la stratégie choisie dans les conditions punition négative et renforcement positif plus punition négative) et les trois gobelets apparaissent à nouveau alignés pour la partie suivante.

A la fin des 100 parties, le score final et la mention « Terminé » apparaissent au

centre l'écran.

Une partie administrative (non accessible aux sujets, excepté à la fin des 100 parties, lors de la vérification de l'emplacement effectif de la balle) contenant pour chaque sujet le tirage au sort de l'emplacement de la balle à chaque partie effectué à l'avance par le programme.

Procédure

Les consignes lues aux sujets ont été intégrées à la procédure, dans la mesure où celles-ci expliquent clairement les étapes des trois protocoles.

Deux expérimentateurs sont nécessaires au bon déroulement de l'expérience (expérimentateur 1 et expérimentateur 2).

L'expérimentateur 1 présente le dilemme de Monty Hall à l'aide de trois gobelets réels de la manière suivante :

« Vous avez devant vous 3 gobelets (*l'expérimentateur 1 montre les 3 gobelets*).

Un jeton rouge se trouve caché sous l'un des gobelets. Vous allez essayer de le trouver. Pour cela, indiquez-moi, sans le soulever, le gobelet dans lequel vous pensez que se trouve ce jeton rouge; même si vous n'avez aucune idée d'où il peut bien se trouver, tentez votre chance, choisissez un gobelet. Allez-y ! (*l'expérimentateur 1 attend que le sujet ait choisi un gobelet*). Bien ! Afin qu'il ne puisse y avoir aucun malentendu, je marque le gobelet que vous venez de choisir en y posant ce jeton blanc dessus (*l'expérimentateur 1 pose le jeton blanc sur le gobelet choisi*) .

Avant de soulever ce gobelet pour voir si vous avez gagné ou perdu, je vais vous aider un peu. Regardez bien, je vais soulever l'un des gobelets que vous n'avez pas choisis et dans lequel ne se trouve pas le jeton rouge (*l'expérimentateur 1 soulève un gobelet*).

Comme vous pouvez le constater, le jeton rouge ne se trouvait pas dans le gobelet que je viens de soulever. Il se trouve donc forcément soit dans celui que vous avez choisi, soit dans l'autre. Avant de soulever le gobelet que vous avez choisi pour voir si vous avez gagné ou perdu, je vais vous donner une deuxième chance : en effet, je vous permets, si vous le souhaitez, de modifier votre choix et de choisir l'autre gobelet ; bien entendu, vous pouvez, si vous le préférez, conserver votre choix initial.

Sachant que votre but est de trouver le jeton rouge, quelle est, selon vous, la stratégie qui vous offre le plus de chances de gagner ? : a) Conserver votre choix initial ; b) Modifier votre choix initial ; c) Cela n'a aucune importance, car vous avez autant de chances de gagner en conservant votre choix qu'en le modifiant. (*l'expérimentateur 1 note la réponse*)

Pouvez-vous justifier votre réponse ? (*l'expérimentateur 1 note la réponse*)

Bien ; voilà ce que vous allez faire à présent. Vous allez jouer 100 fois consécutives sur un logiciel informatique qui reproduit exactement les conditions de ce petit jeu. A chaque partie, une balle sera cachée sous l'un des 3 gobelets, vous devrez en choisir un, puis, une fois ce dernier marqué, un des gobelets que vous n'avez pas choisis et qui ne contiendra pas la balle sera retiré ; vous pourrez alors choisir de conserver ou de modifier votre choix initial et le gobelet correspondant à votre choix de stratégie sera soulevé pour voir si vous avez gagné ou perdu.

Selon la condition expérimentale, l'expérimentateur 1 poursuit :

- *Condition renforcement positif* : « Lorsque votre choix de stratégie (Conserver ou Modifier le choix initial) vous permettra de trouver la balle, vous remporterez un point qui sera ajouté au compteur. Le compteur démarre à 0. »

- *Condition punition négative* : « Lorsque votre choix de stratégie (Conserver ou Modifier le choix initial) ne vous permettra pas de trouver la balle, vous perdrez un point qui sera retiré au compteur. Le compteur démarre à 100. »

- *Condition renforcement positif et punition négative* : « Lorsque votre choix de stratégie (Conserver ou Modifier le choix initial) vous permettra de trouver la balle, vous remporterez un point qui sera ajouté au compteur. Lorsque votre choix de stratégie ne vous permettra pas de trouver la balle, vous perdrez un point qui sera retiré au compteur. Le compteur démarre à 0. »

L'expérimentateur 1 termine la présentation du jeu :

« Le score qui apparaîtra sur le compteur à la fin des 100 parties sera reporté dans un classement dans lequel figureront les résultats de tous les participants à cette expérience. Les 6 premiers de ce classement (dans l'ordre décroissant du nombre de points figurant sur le compteur à la fin des 100 parties) gagneront une somme en euros (le premier et le second recevront 30 euros chacun ; le troisième et le quatrième, 20 euros chacun ; le cinquième et le sixième, 10 euros chacun). Vous avez donc tout intérêt à rester très concentré durant le déroulement des 100 parties.

Vous avez bien compris ? Sinon, je peux vous relire tout ou une partie de la consigne, autant de fois qu'il le faudra, jusqu'à ce que vous ayez bien compris. »

L'expérimentateur 1 donne au sujet le règlement du jeu officiel pour qu'il puisse le lire attentivement. Il le conduit ensuite dans la salle où se trouve l'ordinateur sur lequel se déroule les 100 parties sur logiciel informatique. L'expérimentateur 2 accueille le sujet et lui donne une enveloppe fermée dans laquelle se trouve le tirage au sort de l'emplacement de la balle pour les 100 parties imprimé sur une feuille. Cette enveloppe reste placée à côté de l'ordinateur durant les 100 parties.

L'expérimentateur 2 explique le fonctionnement du logiciel informatique au sujet puis le laisse effectuer les 100 parties librement, au rythme qui lui convient, sans intervenir.

Lorsque les 100 parties sont terminées, le score final apparaît au centre de l'écran.

L'expérimentateur 2 pose alors la question suivante :

« A présent que vous avez effectué les 100 parties, quelle est la stratégie qui vous semble vous offrir le plus de chance de trouver la balle ? a) Conserver le choix initial ; b) Modifier le choix initial ; c) Cela n'a aucune importance car vous avez autant de chances de gagner en conservant qu'en modifiant votre choix initial ». L'expérimentateur 2 note la réponse.

Ensuite, l'expérimentateur 2 ouvre l'enveloppe contenant le tirage au sort de l'emplacement de la balle pour les 100 parties devant le sujet, et ouvre la page de la partie administrative du programme informatique contenant l'emplacement de la balle à chaque partie. Le sujet vérifie avec l'expérimentateur 2 l'exacte correspondance entre les deux tirages.

L'expérimentateur 1 conduit alors le sujet dans la salle où les règles du jeu lui ont été présentées. Il lui pose la question suivante :

« A présent que vous avez effectué les 100 parties et que vous avez pu vérifier la correspondance exacte entre le tirage au sort imprimé de l'emplacement de la balle pour chaque partie et celui de la partie administrative du programme informatique, quelle est la stratégie qui vous semble vous offrir le plus de chance de trouver la balle ? a) Conserver le choix initial ; b) Modifier le choix initial ; c) Cela n'a aucune importance car vous avez autant de chances de gagner en conservant qu'en modifiant votre choix initial. » L'expérimentateur 1 note la réponse.

L'expérimentateur 1 fait signer la feuille d'engagement (stipulant que le participant ne doit rien révéler de l'expérience) et remercie le participant.

RÉSULTATS

Il a été vérifié que le pourcentage de réussite produit par chacune des deux stratégies (« Conserver » ou « Modifier ») lors des 100 parties se trouve dans les limites de confiance des probabilités théoriques (1/3 versus 2/3). Par ailleurs, aucun sujet n'a modifié sa réponse entre les deux dernières questions, c'est-à-dire que les sujets n'ont pas changé leur réponse concernant l'efficacité des stratégies après avoir pu vérifier la correspondance entre le tirage au sort préalable et l'emplacement de la balle qui a été effectivement observé. C'est pourquoi les réponses aux deux dernières questions sont présentées sans distinction dans le graphique n° 3.

Tests statistiques employés

Les comparaisons concernant les nombres moyens de stratégies « Modifier » choisies sont effectuées à l'aide du test t (échantillons indépendants pour les comparaisons inter-groupes et échantillons appariés pour les comparaisons intra-groupes) au seuil $p < .05$.

Les comparaisons concernant les proportions de sujets par catégorie de réponses verbales sont effectuées à l'aide du test z (échantillons appariés) au seuil $p < .05$.

Comparaisons inter-groupes

Réponses verbales à la première question. Aucun sujet, quelle que soit la condition, n'a considéré que « Modifier » était la stratégie optimale. La proportion de sujets considérant que les deux stratégies sont équivalentes à la première question est à peu près la même dans les trois conditions (RP + PN = 86,7 % ; RP = 86,3 % ; PN = 93,3 %). La proportion de sujets considérant que la stratégie « Conserver » était optimale est, bien sûr, complémentaire dans chaque condition. Compte tenu de la taille des effectifs, nous n'avons pas réalisé de comparaison statistique de ces proportions de réponses verbales.

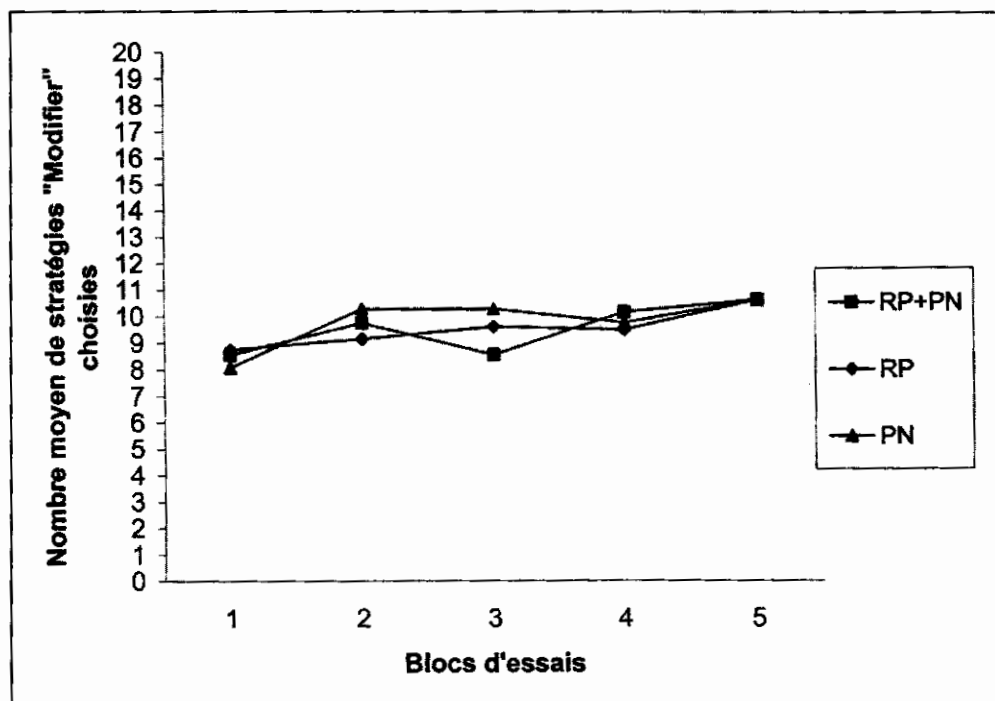
Réponses verbales aux deux dernières questions. La proportion de sujets considérant que la stratégie « Modifier » est optimale aux deux dernières questions est plus importante dans la condition RP + PN (20 %) que dans la condition RP (13,3 %), mais moins importante que dans la condition PN (26,7 %). On observe qu'aucun sujet des trois conditions n'a considéré que la stratégie « Conserver » était optimale aux deux dernières questions. Les proportions de sujets qui ont considéré que les deux stratégies étaient équivalentes sont, bien sûr, complémentaires aux proportions de sujets qui ont considéré que la stratégie « Modifier » était optimale. Compte tenu de la taille des effectifs, nous n'avons pas réalisé de comparaison statistique de ces proportions de réponses verbales.

Tableau 1

| BLOCS | RP+PN | | | RP | | | PN | | |
|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | Moyenne | % | Ecart type | Moyenne | % | Ecart type | Moyenne | % | Ecart type |
| 1 | 8,53 | 42,6 | 4,21 | 8,73 | 43,6 | 5,42 | 8,06 | 40,3 | 4,92 |
| 2 | 9,73 | 48,6 | 3,69 | 9,13 | 45,6 | 3,7 | 10,26 | 51,3 | 4,46 |
| 3 | 8,53 | 42,6 | 4,19 | 9,6 | 48 | 4,19 | 10,26 | 51,3 | 4,73 |
| 4 | 10,13 | 50,6 | 3,89 | 9,46 | 47,3 | 5,21 | 9,73 | 48,6 | 4,67 |
| 5 | 10,6 | 53 | 3,6 | 10,6 | 53 | 4,87 | 10,6 | 53 | 4,64 |
| Total | 47,53 | 47,53 | 15,22 | 47,53 | 47,53 | 19,99 | 48,93 | 48,93 | 18,69 |

Nombre moyen de stratégies « Modifier » choisies au cours des 100 essais par blocs de 20 essais en fonction du type de procédure d'apprentissage.

Nombre moyen de stratégies « Modifier » choisies. On observe un nombre moyen de stratégies « Modifier » choisies au cours des 100 essais égal pour les sujets des conditions RP + PN et RP (47,53). On observe un nombre moyen de stratégies « Modifier » choisies au cours des 100 essais légèrement plus important dans la condition PN (48,93) que dans les conditions RP et RP + PN (47,53). Ces différences ne sont cependant pas significatives (respectivement $t = 0,249$ et $t = 0,352$; $p > .05$).



Graphique 1. Evolution du nombre moyen de stratégies « Modifier » choisies au cours des 100 essais par blocs de 20 essais en fonction du type de procédure d'apprentissage

On observe que les courbes des trois conditions suivent une évolution similaire, c'est-à-dire, globalement ascendante, mais faiblement. Le nombre moyen de stratégies « Modifier » choisies par les sujets des trois conditions évolue d'environ 8,5 pour les 20 premiers essais à environ 10,5 pour les 20 derniers essais.

Par ailleurs, aucune différence significative n'apparaît entre les trois conditions à aucun moment de l'expérience. On observe que le nombre moyen de stratégies « Modifier » choisies au cours des 20 premiers essais est légèrement plus important dans la condition RP (8,73) que dans la condition RP + PN (8,53). Cependant, cette

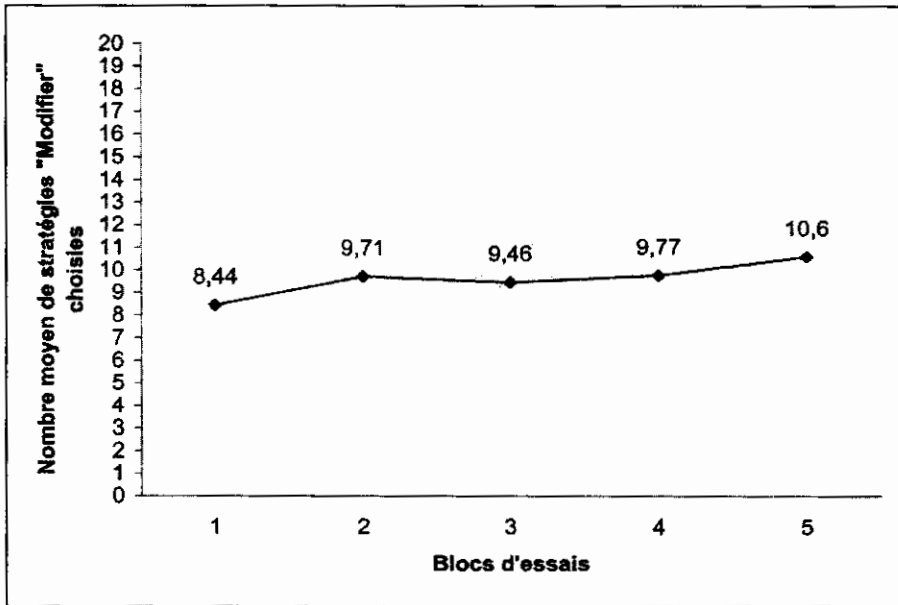
différence n'est pas significative ($t = 0,023$; $p > .05$)

On observe que le nombre moyen de stratégies « Modifier » choisies au cours des 20 premiers essais est légèrement plus important dans la condition RP (8,73) que dans la condition PN (8,06). Cependant, cette différence n'est pas significative ($t = 0,362$; $p > .05$).

On observe que le nombre moyen de stratégies « Modifier » choisies au cours des 20 premiers essais est légèrement plus important dans la condition RP + PN (8,53) que dans la condition PN (8,06). Cependant, cette différence n'est pas significative ($t = 0,405$; $p > .05$).

On observe que le nombre moyen de stratégies « Modifier » choisies au cours des 20 derniers essais est le même dans les trois conditions (10,6).

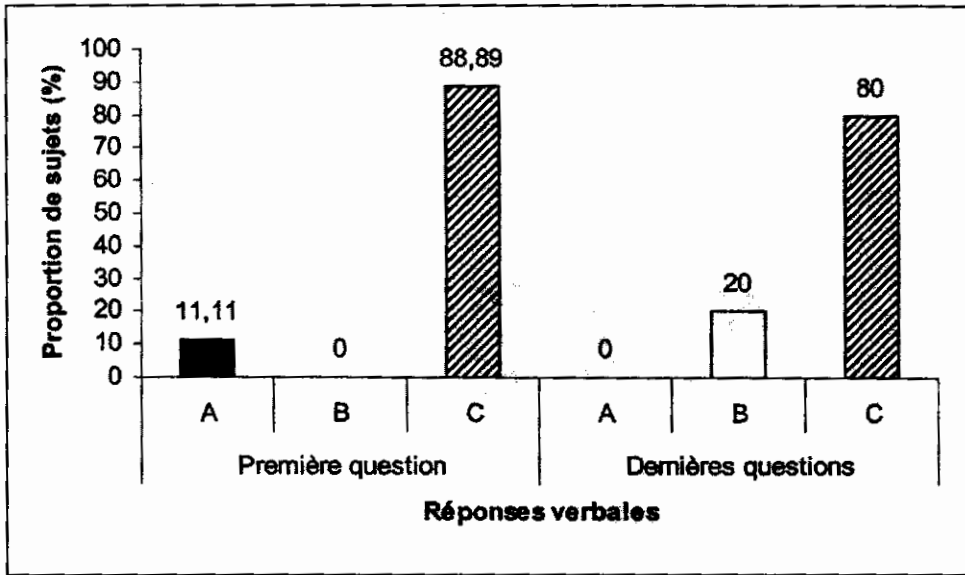
Comparaisons intra-groupe : trois conditions regroupées



Graphique 2. Evolution du nombre moyen de stratégies « Modifier » choisies au cours des 100 essais par blocs de 20 essais pour l'ensemble des sujets des trois conditions

Nombre moyen de stratégies « Modifier » choisies. On observe que la courbe de l'évolution du nombre moyen de stratégies « Modifier » choisies est ascendante, mais faiblement. Seul le troisième bloc d'essais montre un léger recul du nombre moyen de stratégies « Modifier ».

On observe que le nombre moyen de stratégies « Modifier » choisies au cours des 20 derniers essais (10,6. Ecart type = 4,30) est plus important que le nombre moyen de stratégies « Modifier » choisies au cours des 20 premiers essais (8,44. Ecart type = 4,77). Cette différence est significative ($t = 3,544$; $p < .05$).



Graphique 3. Proportion de sujets de l'ensemble des trois conditions par catégorie de réponses verbales (A: Stratégie « Conserver » considérée optimale; B: Stratégie « Modifier » considérée optimale; C: Stratégie « Conserver » et Stratégie « Modifier » considérées équivalentes).

Réponses verbales. Première question :

On observe que la proportion de sujets considérant que la stratégie « Conserver » est optimale (11,11 %) est plus importante que la proportion de sujets considérant que la stratégie « Modifier » est optimale (0 %). Cette différence est significative ($z = 2,29$; $p < .05$).

On observe que la proportion de sujets considérant que les deux stratégies sont équivalentes (88,89 %) est plus importante que la proportion de sujets considérant que la stratégie « Conserver » est optimale (11,11 %). Cette différence est significative ($z = 7,37$; $p < .05$).

On observe que la proportion de sujets considérant que les deux stratégies sont équivalentes (88,89 %) est plus importante que la proportion de sujets considérant que

la stratégie « Modifier » est optimale (0 %). Cette différence est significative ($z = 8,45$; $p < .05$).

Deux dernières questions :

On observe que la proportion de sujets considérant que la stratégie « Modifier » est optimale (20 %) est plus importante que la proportion de sujets considérant que la stratégie « Conserver » est optimale (0 %). Cette différence est significative ($z = 3,16 > 1,96$; $p < .05$).

On observe que la proportion de sujets considérant que les deux stratégies sont équivalentes (80 %) est plus importante que la proportion de sujets considérant que la stratégie « Conserver » est optimale (0 %). Cette différence est significative ($z = 7,74$; $p < .05$).

On observe que la proportion de sujets considérant que les deux stratégies sont équivalentes (80 %) est plus importante que la proportion de sujets considérant que la stratégie « Modifier » est optimale (20 %). Cette différence est significative ($z = 5,69$; $p < .05$).

Evolution des réponses verbales :

On observe que la proportion de sujets ayant considéré que la stratégie « Modifier » était optimale est plus importante aux deux dernières questions (20 %) qu'à la première question (0 %). Cette différence est significative ($z = 3,16$; $p < .05$).

On observe que la proportion de sujets ayant considéré que la stratégie « Conserver » était optimale est moins importante aux deux dernières questions (0 %) qu'à la première question (11,11 %). Cette différence est significative ($z = 2,30$; $p < .05$).

On observe que la proportion de sujets ayant considéré que les deux stratégies étaient équivalentes est moins importante aux deux dernières questions (80 %) qu'à la première question (88,89 %). Cependant, cette différence n'est pas significative ($z = 1,16$; $p > .05$).

DISCUSSION

Comparaisons inter-groupes

Les résultats obtenus en comparaison inter-groupes semblent indiquer qu'il faille rejeter l'hypothèse selon laquelle l'association entre le renforcement positif et la punition négative accroît l'efficacité de l'exposition concrète aux contingences dans l'apprentissage du comportement optimal par rapport au renforcement positif ou à la punition négative

employés isolément. La procédure de renforcement positif n'a pas non plus été plus efficace que la procédure de punition négative, contrairement à l'hypothèse que nous avons formulée sur la base des travaux de Dinsmoor (1983 – cité par Catania, 1994). Aucune procédure ne s'est donc avérée plus efficace qu'une autre.

Comparaisons intra-groupe

Pour l'ensemble des trois conditions, on observe que l'exposition concrète aux contingences a eu une certaine efficacité dans l'apprentissage du comportement optimal, puisque le nombre moyen de stratégies optimales choisies augmente significativement au fur et à mesure des essais (en moyenne, environ deux choix supplémentaires de la stratégie optimale entre le premier bloc et le dernier bloc de 20 essais). Il faut cependant noter que cette augmentation est faible, et que les taux observés restent proches de 50 %.

Au niveau des réponses verbales, on constate qu'à l'énoncé du problème, aucun sujet, quelle que soit la condition expérimentale, n'a considéré que la stratégie « Modifier » était optimale, et qu'une majorité significative de sujets a considéré que les deux stratégies étaient équivalentes. Ce résultat est concordant avec l'ensemble des études expérimentales concernant le dilemme de Monty Hall.

Après les 100 essais, la répartition des sujets en fonction de leur réponse verbale aux deux dernières questions est différente, puisque la proportion de sujets considérant que « Conserver » est la stratégie optimale diminue significativement et que la proportion de sujets qui considèrent que « Modifier » est la stratégie optimale augmente significativement. On peut donc conclure à une efficacité réelle, bien que faible, de l'exposition concrète aux contingences dans l'apprentissage du choix optimal.

Cependant, la proportion de sujets considérant que les deux stratégies sont équivalentes reste significativement et très largement majoritaire. Cela indique que l'exposition aux contingences n'a pas changé la suprématie du comportement non-optimal verbalement formulé. Seul le choix non optimal de la stratégie « Conserver » a disparu.

DISCUSSION GÉNÉRALE

On constate, en début d'apprentissage, un nombre moyen de comportements optimaux supérieur aux taux observés dans de nombreuses études ayant précédemment employé des séries d'essais renforcés dans le dilemme de Monty Hall. Dans la présente étude, un taux proche de 50 % est obtenu dès le premier bloc de 20 essais, alors que, dans les études précédentes, les taux sont proches de 20 % à l'exposé du problème et en début

d'apprentissage (Granberg et Brown, 1995 ; Granberg et Dorr, 1998 ; Friedman, 1998 ; Mueser et Granberg, 1999 ; Slembeck et Tyran, 2002 ; Bown, Read et Summers, 2002 ; Tubau et Alonso, 2003). Bien que les proportions de choix calculées dans ces études ne soient pas toujours faciles à comparer à celles de la présente étude, puisque le nombre d'essais n'est pas toujours le même, il semble bien que les sujets aient eu ici tendance à choisir chacune des deux stratégies avec une fréquence très proche l'une de l'autre tout à long des 100 essais, alors qu'il avait été plutôt observé dans d'autres études une tendance forte à conserver le choix initial puis une augmentation du choix de la stratégie « Modifier ».

Des hypothèses ont déjà été avancées pour expliquer la contradiction entre le jugement des sujets à l'exposition du dilemme de Monty Hall (c'est-à-dire, considérer les deux stratégies comme étant équivalentes) et le choix préférentiel de la stratégie « Conserver ». Ainsi, Mueser et Granberg (1999) et Slembeck et Tyran (2002), qui s'appuient notamment sur le concept de raisonnement contre-factuel, expliquent qu'il est plus regrettable pour les sujets de ne pas gagner en modifiant que de ne pas gagner en conservant. Le regret est anticipé comme étant comparativement plus fort quand les sujets pensent qu'ils avaient trouvé où était la balle avec leur premier choix puis ont modifié par rapport à quand ils n'avaient pas trouvé où était la balle puis ont conservé. Les sujets se sentiraient engagés dans leur premier choix car ils voudraient éviter le regret.

La notion d'engagement a également été développée dans le cadre de la théorie de l'engagement de Kiesler (1971-cité par Beauvois et Joule, 1987, 1998 et par Cialdini, 1990-) : un individu aurait tendance à émettre des comportements allant dans le même sens que ceux précédemment émis. Dans le dilemme de Monty Hall, le sujet doit prendre deux décisions consécutives, dont la deuxième remet en cause la première. En effet, il doit d'abord choisir un gobelet parmi trois pour trouver la balle. Ensuite, il doit choisir de conserver ou de modifier son choix initial. Dans le cadre de la théorie de l'engagement, on peut avancer l'hypothèse selon laquelle un effet d'engagement serait impliqué dans la persistance pour le choix initial. Une première décision étant prise pour le choix d'un gobelet, on peut en effet s'attendre à ce que le sujet soit engagé dans cette décision et préfère conserver le gobelet désigné au départ. L'effet d'engagement pourrait expliquer en partie pourquoi les sujets, alors qu'ils considèrent les deux stratégies équivalentes, ont tendance à choisir la stratégie « Conserver » en début d'apprentissage dans la plupart des études expérimentales.

N'ayant pas obtenu des résultats similaires dans la présente étude, une nouvelle hypothèse peut être avancée afin d'expliquer, d'une part, pourquoi on obtient un taux de comportements optimaux en début d'apprentissage supérieur aux études précédentes et, d'autre part, pourquoi l'augmentation du nombre moyen de comportements optimaux est si faible, même après l'exposition concrète à 100 essais.

La littérature montre que le suivi d'une règle a tendance à améliorer l'efficacité de l'exposition aux contingences qui vont dans le même sens que cette règle (Rosenfarb, Newland, Brannon et Howey, 1992).

Dans la présente étude, l'opinion des participants concernant l'efficacité des stratégies a été relevée avant l'exposition aux contingences, et cette opinion était exprimée au travers d'une réponse à choix multiples à une question fermée. Dans les études précédentes, cette opinion était relevée au moyen d'un questionnaire post-expérimental. Il est très possible que le fait d'avoir proposé directement les trois solutions avant la session d'apprentissage ait réduit la probabilité d'apparition du choix de la stratégie « Conserver » en amenant les participants à réfléchir à toutes les possibilités qui s'offraient à eux en matière de jugement de probabilité, et en les conduisant à formuler une règle qu'ils ont immédiatement appliquée.

En revanche, il a été montré que le suivi d'une règle a tendance à entraîner chez un sujet une insensibilité aux contingences contradictoires avec cette règle, y compris lorsque celle-ci est formulée par le sujet lui-même (Rosenfarb, Newland, Brannon et Howey, 1992). On peut ainsi avancer l'hypothèse selon laquelle les participants ont émis un comportement cohérent avec la règle que nous les avons amenés à formuler, c'est-à-dire qu'ils ont choisi chacune des deux stratégies en proportion équivalente, et ont été peu sensibles aux contingences qui contredisaient cette règle, c'est-à-dire au renforcement deux fois plus fréquent de la stratégie « Modifier ».

Par ailleurs, une autre explication possible de la difficulté qu'un individu, considérant que les deux choix sont équivalents, rencontre à apprendre le comportement optimal réside dans la faible différence qui existe entre des taux de renforcement de $2/3$ versus $1/3$ par rapport au taux de $1/2$ pour chaque stratégie. Comme l'indiquent Slembeck et Tyran (2002), même un individu initié aux tests statistiques rejeterait difficilement l'hypothèse des stratégies équivalentes, car, d'après un test binomial, pour 10 essais, il lui faudrait observer 7 succès de la stratégie « Modifier », 14 succès pour 20 essais ou encore 27 succès après 40 essais. Ces taux sont proches de 50 %, et il apparaît bien difficile de conclure à une supériorité de la stratégie « Modifier » sans avoir recours à un calcul mathématique.

En conclusion, il apparaît que de nouvelles recherches concernant l'apprentissage du comportement optimal dans le dilemme de Monty Hall peuvent s'orienter vers la modification des règles verbales suivies par les individus. En effet, il semble que les difficultés que les individus rencontrent à corriger leurs erreurs systématiques de jugement en situation d'incertitude pourraient désormais être appréhendées sous l'angle de l'insensibilité aux contingences due au suivi de règles verbales. Cette perspective pourrait s'avérer féconde dans la mesure où, jusqu'à présent, aucune étude ayant employé des séries d'essais renforcés pour l'apprentissage du comportement optimal dans le dilemme de Monty Hall n'a obtenu une compréhension des règles de probabilités qui permettent

d'aboutir à la solution correcte.

BIBLIOGRAPHIE

- Aaron, E. et Spivey, M. (1998). Frequency vs Probability Formats : Framing the Three Doors problem. In Gernsbacher, M. A. et Derry, S. J. Proceedings of the Twentieth Annual Conference of the Cognitive Science Society (13-18). Mahwah : Lawrence Erlbaum Associates.
- Bown, N. J. , Read, D. et Summers, B. (2002). The Lure of Choice. *Working paper at Department of Operational Research London School of Economics and Political Science.*
- Catania, A., C. et Cutts, D.(1963). Experimental Control of Superstitious Responding in Humans. *Journal of the Experimental Analysis of Behavior*, 6, 203-208
- Catania, A. C. (1994). Information, Cognition and Behavior : telling what they want to hear. Book review of Giere on scientific cognition. Disponible à l'adresse internet : <http://www.cogsci.ecs.soton.ac.uk/psychology/raw/1994.volume.5/psyc.94.5.39.scientific-cognition.5.catania>
- Chomsky, N. (1967). Appendice. In H. M. Lenneberg (ed.), *Biological foundations of language*. New York, London, Sidney : John Wiley and Sons. Version française : (1969). La nature formelle du langage. Paris : Editions du Seuil.
- Cialdini, R. (1990). Influence et Manipulation. Paris: First
- Fantino, E. (1998). Judgment and Decision Making: Behavioral Approaches. *The Behavior Analyst*, 21, 203-218.
- Franco-Watkins, A., Derks, P. et Dougherty, M. (2003). Reasoning in the Monty Hall Problem: Examining Choice Behaviour and Probability Judgments. *Thinking and Reasoning*, 1, 67-90.
- Friedman, D. (1998). Monty Hall's Three Doors : Construction and Deconstruction of a Choice Anomaly. *American Economic Review*, 88, 933-946.
- Goodie, A. S., et Fantino, E. (1996). Base Rates versus Sample Accuracy : Competition for Control in Human Matching to Sample. *Journal of the Experimental Analysis of Behavior*, 71, 155-169.
- Granberg, D. (1999). Cross-Cultural Comparison of Responses to the Monty Hall Dilemma. *Social Behavior and Personality*, 27, 431-438.
- Granberg, D. et Brown, T. A. (1995). The Monty Hall Dilemma. *Personality and Social Psychology Bulletin*, 21, 711-723.
- Grandberg, D. et Dorr, N. (1998). Further Exploration of Two-Stage decision Making in the Monty Hall Dilemma. *American Journal of Psychology*, 111, 561-576.
- Joule, R.-V. et Beauvois, J.-L. (1987). *Petit Traité de Manipulation à l'Usage des Honnêtes Gens*. Grenoble : Presses Universitaires de Grenoble.
- Joule, R.-V. et Beauvois, J.-L. (1998). *La Soumission Librement Consentie*. Paris : Presses Universitaires de France.
- Kahneman, D. et Tversky, A. (1973). On the Psychology of Prediction. *Psychological Review*, 80, 237-251.
- Kahneman, D., Slovic, P., et Tversky, A. (1982). *Jugement Under Uncertainty: Heuristics and Biases*. Cambridge : Cambridge University Press.
- Krauss, S. et Wang, X. T. (2003). The Psychology of the Monty Hall Problem: Discovering Psychological Mechanisms for Solving a Tenacious Brain Teaser. *Journal of Experimental Psychology-General*, Vol. 132, 1, 3-22.

- Mueser, P. R. et Granberg, D. (1999). The Monty Hall Dilemma Revisited : Understanding the Interaction of Problem Definition and Decision Making. *Economics Working Paper Archive at WUSTL*.
- Palacios-Huerta, I. (2003). Learning to Open Monty Hall's Doors. *Experimental Economics*, Vol. 6(3), 235-251.
- Piattelli Palmarini, M. (1993). L'illusion de Saperere. Milan : Arnoldo Mondadori Editore. Version française (1995). *La Réforme du Jugement ou Comment ne Plus Se Tromper*. Paris : Odile Jacob.
- Rosenfarb, I. S., Newland, M. C., Brannon, S. E. et Howey, D. S. (1992). Effects of Self-Generated Rules on the Development of Schedule-controlled Behavior. *Journal of the Experimental Analysis of Behavior*, 58, 107-121.
- Scholiers, D., Villatte, M. et Freixa i Baqué, E. (2003). Comportement Non Optimal Versus Illusion Cognitive. *Acta Comportamentalia*, 11, 5-17.
- Shier D. (2000). Can Human Rationality Be Defended a Priori ? *Behavior and Philosophy*, 28, 67-81.
- Slembeck, T. et Tyran, J. R. (2002). Do Institutions Promote Rationality ? An Experimental Study of the Three-Door Anomaly. *A paraître in Journal of Economic Behavior and Organization*.
- Stich, S. (1997). Cognitive Penetrability. *Mind & Language*, 12, 297-326.
- Stich, S. (1998). Rethinking Cognition. *Mind & Language*, 13, 499-512.
- Stich, S. (1999). *Handbook of Epistemology*. Dordrecht: Matti Sintonen.
- Stolarz-Fantino, S., Fantino, E., Zizzo, D.J. et Wen, J. (2003). The Conjunction Fallacy : New Evidences for Robustness. *American Journal of Psychology*, 116,15-34.
- Tubau, E. et Alonso, D. (2003). Overcoming Illusory Inferences in a Probabilistic Counterintuitive Problem : The Role of Explicit Representations. *Memory and Cognition*, 31, 596-604.
- Wason, P., et Johnson-Laird, P. (1970) A Conflict Between Selecting and Eliminating Information in an Inferential Task, *British Journal of Psychology*, 61, 509-515.

RÉSUMÉ

L'étude a pour objet le dilemme de Monty Hall, une situation de jugement de probabilités dans laquelle la très grande majorité des individus émet un comportement de choix non-optimal.

Nous formulons l'hypothèse selon laquelle l'exposition répétée aux conséquences du comportement de choix va permettre l'apprentissage du comportement optimal. Trois conditions constituent l'expérience : une condition consistant en une procédure de renforcement positif, une condition consistant en une procédure de punition négative, et une condition consistant en une procédure de renforcement positif et de punition négative. L'ajout d'une valeur aversive en conséquence du choix non-optimal, ainsi que l'extension de la série d'essais à 100, devraient permettre une amélioration de l'efficacité de l'apprentissage par rapport aux études ayant précédemment employé des séries d'essais renforcés dans le dilemme de Monty Hall.

Les résultats montrent que le taux de comportement optimal augmente avec la série d'essais, mais reste proche de 50 %. Aucune des trois conditions d'apprentissage ne se révèle plus efficace qu'une autre pour l'acquisition du choix optimal. L'hypothèse selon laquelle le comportement de choix des sujets serait gouverné par des règles verbales les rendant insensibles aux contingences changeantes est formulée.

Mots clés : rationalité, jugement en situation d'incertitude, dilemme de Monty Hall, comportement non-optimal, illusion cognitive, règles verbales, apprentissage.

ABSTRACT

Judgement under uncertainty has generated many research in Cognitive Sciences where the erroneous judgment is perceived as the product of mental modules unsuited to the situation and in Experimental Analysis of Behavior, where it is apprehended as a non-optimal behavior acquired by learning.

The experiment consists of a game, the Monty Hall Dilemma, in which 3 goblets are presented turned over to a subject who must find a ball hidden under one of them. After a first choice, an empty and nonselected goblet is withdrawn. The subject can then stick with his initial choice or switch to the other remaining goblet. The "Switch" strategy and the "Stick" strategy have, respectively, a profit probability of $2/3$ and $1/3$, but, in great majority, the subjects consider that the two strategies are equivalent, what constitutes a non optimal behavior.

We formulate the hypothesis according to which repeated exposure to choice consequences will allow learning of the optimal behavior. Three conditions constitute the experiment: a condition consisting of a positive reinforcement procedure, a condition consisting of a negative punishment procedure, and a condition consisting of a positive reinforcement and negative punishment procedure. The addition of an aversive value when the choice of strategy does not make it possible to find the ball, and the extension of the series of tries to 100 should allow an improvement of the learning effectiveness compared to the studies having previously employed series of reinforced tries in the Monty Hall Dilemma.

Results show that optimal behavior rate increases with series of tests, but remains close to 50 %. No learning procedure appears more effective than another for acquisition of the choice for "Switch" strategy. The hypothesis according to which subjects choice behavior would be controlled by verbal rules making them insensitive with changing contingencies is formulated.

Key words: rationality, judgement under uncertainty, Monty Hall dilemma, non-optimal behavior, cognitive illusion, verbal rules, learning.

ABSTRACT

Judgement under uncertainty has generated many research in Cognitive Sciences where the erroneous judgment is perceived as the product of mental modules unsuited to the situation and in Experimental Analysis of Behavior, where it is apprehended as a non-optimal behavior acquired by learning.

The experiment consists of a game, the Monty Hall Dilemma, in which 3 goblets are presented turned over to a subject who must find a ball hidden under one of them. After a first choice, an empty and nonselected goblet is withdrawn. The subject can then stick with his initial choice or switch to the other remaining goblet. The "Switch" strategy and the "Stick" strategy have, respectively, a profit probability of $2/3$ and $1/3$, but, in great majority, the subjects consider that the two strategies are equivalent, what constitutes a non optimal behavior.

We formulate the hypothesis according to which repeated exposure to choice consequences will allow learning of the optimal behavior. Three conditions constitute the experiment: a condition consisting of a positive reinforcement procedure, a condition consisting of a negative punishment procedure, and a condition consisting of a positive reinforcement and negative punishment procedure. The addition of an aversive value when the choice of strategy does not make it possible to find the ball, and the extension of the series of tries to 100 should allow an improvement of the learning effectiveness compared to the studies having previously employed series of reinforced tries in the Monty Hall Dilemma.

Results show that optimal behavior rate increases with series of tests, but remains close to 50 %. No learning procedure appears more effective than another for acquisition of the choice for "Switch" strategy. The hypothesis according to which subjects choice behavior would be controlled by verbal rules making them insensitive with changing contingencies is formulated.

Key words: rationality, judgement under uncertainty, Monty Hall dilemma, non-optimal behavior, cognitive illusion, verbal rules, learning.