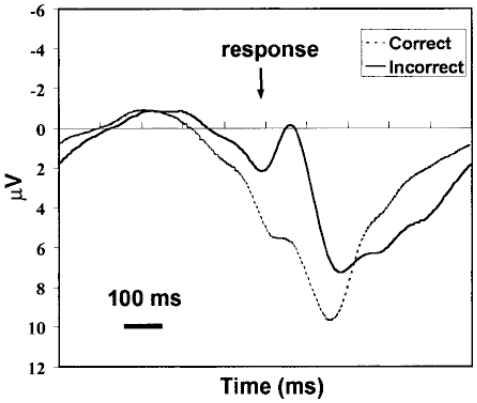
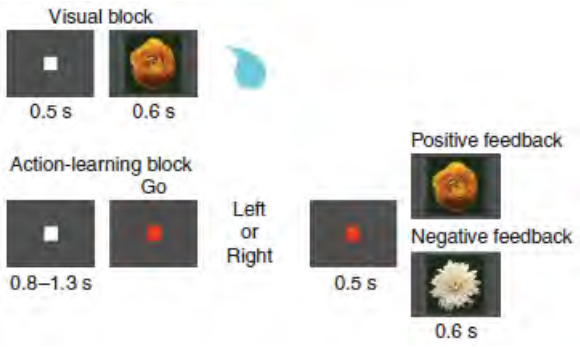
ミン細胞からの予測誤差信号がこれらの脳領域に伝えられることで、報酬量を最大化するような行動が学習されていくと考えるのである。さまざまな強化学習モデルが考えられてきたが、行動の選択肢が複数あるとき、それぞれの選択肢について、それを選んだとしたら、その後の将来にわたって、どれだけの報酬獲得が期待されるかという見積もり（行動価値）を、線条体の神経細胞は主に表現していることを、Samejima ら{Samejima, 2005 #3278}は明らかにした。彼らは、ジョイスティックを右か左に倒すと、ブロック内では一定の確率で報酬のジュースが多めに貰える自由選択課題（上図 A）をサルに訓練した。右にジョイスティックを倒したときに多めの報酬が貰える確率（報酬確率）と左にジョイスティックを倒したときの報酬確率を独立に操作できるのがこの課題の特長である。実際、左報酬確率と右報酬確率の組み合わせは、(90, 50), (50, 90), (50, 10), (10, 50), (50, 50)の５通りを用い（上図 B)、ブロック間でランダムにその条件を入れ替えた。その結果、線条体の多くの神経細胞は、右左どちらにジョイスティックを倒したかに関わらず、ジョイスティックを右もしくは左に倒す行動の価値を表現していた。このことは、行動の選択肢が複数存在する行動的意思決定の場面において、それぞれの行動価値が線条体で表現され、それらを比較した結果として、行動が選択されるということを示唆している。

強化学習モデルが、前頭前野や帯状皮質前部における神経細胞活動を説明するのにも有効であることもわかってきた。Holroyd と Coles {Holroyd, 2002 #2620}は、間違いをおかしたときに帯状皮質前部から記録される陰性（-側）の事象関連電位（右図）は、ドーパミン細胞からの入力に依存しているとする仮説を、数理モデルと併せて提案した。実際、





Matsumoto ら{Matsumoto, 2007 #3653}は、帯状皮質前部の少し前の前頭前野内側部では、行動価値の予測誤差が表現されていることを明らかにした。彼らは、左右のレバー押し学習課題（左図）をサルに訓練し、課題遂行中の神経細胞活動を前頭前野内側部から記録した。数試行からなるブロック内では、正解のレバーは固定していたが、ブロックが切り替わる毎に、正解のレバーをランダムに設定し直した。そして正解と不正解とを視覚的フィードバック刺激の提示によってサルに教えた。このとき、フィードバック刺激に対する細胞応答を解析した

ところ、行動価値の、正負の予測誤差を別々の神経細胞群が表現していた。

前頭前野を含む前頭葉と線条体は、いずれもドーパミン細胞の解剖学的投射を受けているだけでなく、視床を介して情報処理ループを形成していることが分かっている{Alexander, 1986 #2498}。このループ構造がどんな役割を果たしているかの詳細が、強化学習モデルの実現メカニズムの解明と共に、今後明らかにされていくであろう。

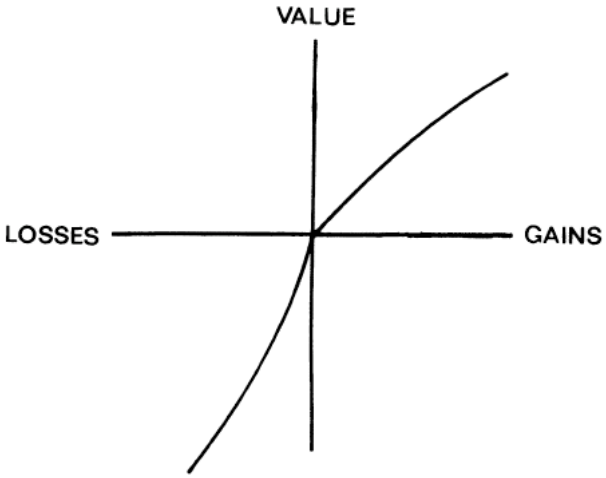
このように、強化学習モデルを援用することにより、行動的意思決定に関わる脳内の情報処理システムの演算処理の理解が急速に進んできた。しかし、学習行動を説明するモデルは一般に複数可能であると思われるので、より真実に近いモデルを見出すためには、モデル間で比較して、最適なモデルをさらに追究していくことが必要である{Daw, 2004 #2978}。さらに、行動選択の脳内過程がモデル通りに実現されているかどうかを検討するには、モデル内部の変数の脳内表現を何らかの方法で直接に操作し、観察可能な行動指標の変化に反映させる必要があると思われる。

（２）神経経済学のレビュー～神経経済学の勃興と現在～

①背景

従来の経済学では、人間は利益を最大化するために合理的な判断をすることが前提とされてきた{Von Neumann, 1944 #3252}。しかし実際は、人間は非合理的な判断、意思決定をすることが多い。例えば、健康のためにはアルコール類は控えた方がよいと分かっていたとしても、私たちはついついビールを飲み過ぎたりしてしまう。

従来の経済学は、最も合理的かつ経済的な判断を前提としており、そのように意思決定すべきという規範を示していることから、規範的経済理論と呼ばれる。それに対して、新しい経済理論＝行動経済学は、必ずしも合理的／経済的でなかったとしても、人は実際にどのように判断、意思決定をするかという心理学的／生物学的要因を考慮しており、記述的経済理論であると言われる。記述的経済理論の代表的なものに、Kahneman と Tversky によるプロスペクト理論がある{Kahneman, 1979 #2963}。この理論が経済学に与えた影響は大きく、その功績により 2002 年のノーベル経済学賞が Kahneman に授与された。プロスペクト理論には、３つの大きな特徴がある。感応度逓減性、参照点依存性、そして損失忌避である。これらの特徴は、Kahneman と Tversky が描いた価値関数曲線（下図）に端的に表れている。



一つめの感応度逓減性とは、客観的な価値の大きさと、主観的な価値の大きさ(＝効用)とは比例関係にはなく、客観的な価値が大きくなるほど、価値と効用との間の変化の割合は減少するという効用の性質のことである。この点は、効用についてのベルヌーイの古典的な考えや、ウェーバー＝フェヒナーの心理物理学的な観点と一致している。価値関数曲線においては、原点から離れるほど、曲線が平坦に近づいているところに表れている。

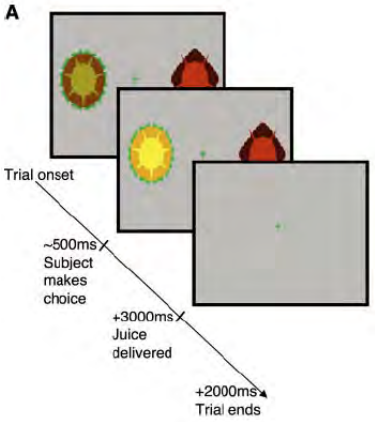
二つ目の参照点依存性とは、価値がゼロである事態が、効用もゼロであることを意味せず、コンテキストによってその効用は変化するという効用の性質である。多くの場合、現状（Status quo）が参照点となることが多い。例えば、貯蓄額がゼロである事態が、価値関数曲線の原点を意味するのではなく、現在の貯蓄額が参照点となり、貯蓄額の増加が正の効用を持ち、貯蓄額の減少が、負の効用を持つ。価値関数曲線においては、原点として参照点が示されており、その参照点は、コンテキストによって変化するということである。

三つ目の損失忌避とは、同じ大きさの価値の増加と減少とを比較すると、減少の方が、インパクトが大きいという効用の性質のことである。例えば、投げたコインが表か裏かによって、同じ額を貰えたり失ったりする賭けゲームを考えたとき、多くの人はこのゲームに参加しないという選択をする。この傾向は、損失忌避から説明される。価値関数曲線においては、正の領域よりも負の領域の方が、曲線の傾きが急峻になっているところに表れている。

行動経済学では、直感システムと論理システムという二つのプロセスを想定し、直感システムによる意思決定の非合理的なバイアスを考慮した。(上図の価値関数は、その一つの表れである。)これにより、多くのジレンマやパラドックスが解消された。しかし直感システムの存在は、行動実験・解析から導かれた理論構築物であって、それをもって行動を説明するというのは一種の循環論法である。伝統的経済学者はここを問題にし、これら二つのシステムが客観的に存在することが明らかでないことを批判した。これに答えるためのツールとして脚光を浴びたのが、機能的磁気共鳴画像撮影 fMRI である。意思決定における直感システムと論理システムの実体を脳に求め、それらを分離することが期待されている。ここに行動経済学と神経科学とが蜜月を迎え、神経経済学が勃興したわけである。神経経済学の発展により、従来の単純で合理的な人間像をもとにした経済学モデルから、脳生理学的現象を基礎にした経済学モデルへの革命が起こるかも知れない。

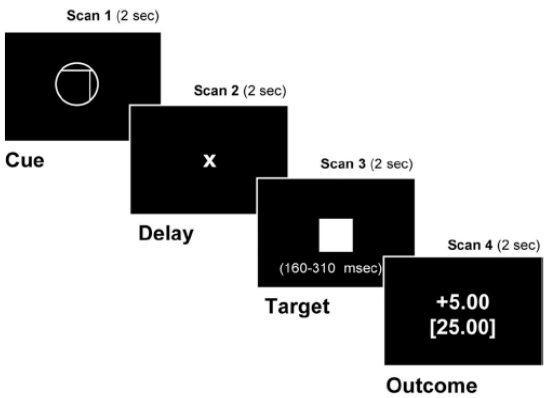
②効用

（１）で述べたように、意思決定の脳内機構については、主にサルなどの動物を用いて調べられてきた。20 世紀の終わり頃になって fMRI や陽電子放射断層撮影法（PET）といった、ヒトの脳活動を画像化する技術が開発されたことにより、ヒトの意思決定の脳内機構を調べることも可能になった。そして、報酬や行動の価値を表現する脳内機構は、動物とヒトとの間で、かなり似通っていることが明らかになってきた。O’Doherty ら{O’Doherty, 2004 #2997}は、二つの図形のうち、より高い確率で美味しいジュースが貰える方を選ぶことを学習する図形選択学習課題（左図）を被験者に行って貰っているときの脳活動を、fMRI を用いて調べた。その結果、動物でも、選択行動の価値に関わっていることが分かっている線条体で、活動が高まっていることが明らかになった。

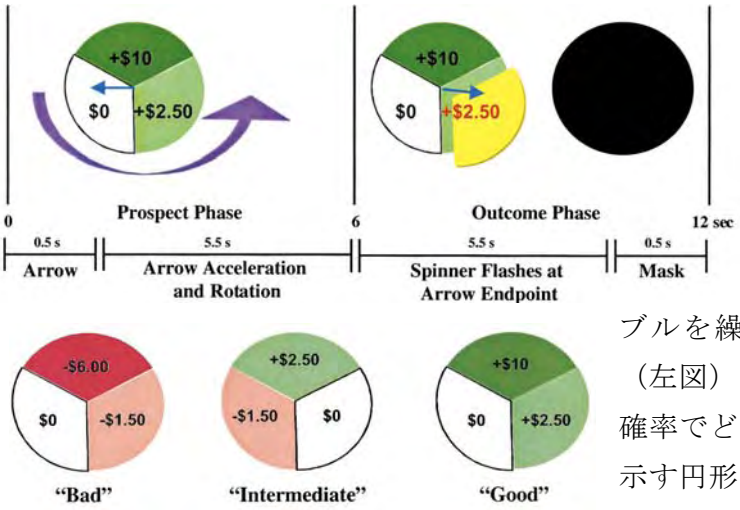


同じような領域が、金銭の価値も表現していることが明らかになっている。意思決定研究（特に行動的意思決定）において注目されたのは、行動が導く結果の価値についての見積もりつまり行動価値であったが、結果の価値自体、そもそも主観的であり、しばしば効用と呼ばれる。効用は従来の経済学においてもプロスペクト理論においても、基本的には得られた際の結果の価値とその得られる確率との積（期待効用）に依存している。

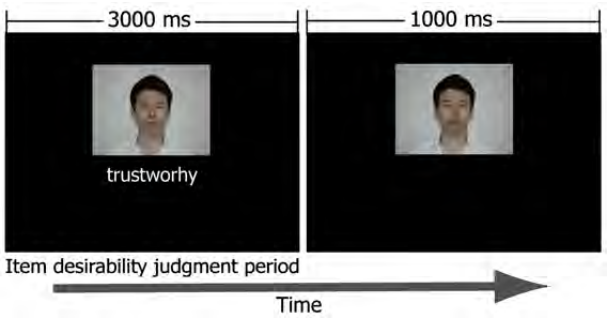
Knutson ら{Knutson, 2005 #3799}は、結果の価値と確率とが脳内でどのように統合されているかを、fMRI を用いて調べた。彼らが用いた課題（右図）では、被験者が画面を見ていると、金銭報酬の額と確率とを示す図形が呈示される。その後、正方形のターゲットが呈示されたとき、できるだけ早くボタンを押せばよいだけの簡単な課題であるが、少しでも反応が遅いと、その金銭報酬を得ることはできない。結果も同じ画面上に呈示される。この課題を遂行しているとき、金銭の額、つまり価値を主に表現したのは、左右の側坐核であり、金銭報酬の確率も含めて表現したのは、前頭前野内側部であった。いずれもドーパミン細胞の投射を強く受ける部位であることから、これらの結果は、ドーパミン細胞の入力が前頭前野や線条体で処理されることによって、期待効用が形成されていることを示唆している。



効用は、それがどのような場面で問題となるかによって変わるので、場面に応じて、経験効用、予期効用、決定効用に分けられる{Kahneman, 1990 #3270}。経験効用とは、実際にある結果を得たことに基づく効用、予期効用とは、ある結果を予期することに基づく効用、そして決定効用とは、実際の行動的意思決定に使われる効用である。３つの効用の間の関係については、現在もまだ明らかではないが、Breiterら{Breiter, 2001 #2714}は、お金の獲得と喪失について、それを予期したときと経験したときの脳活動を、fMRIを用いて調べた。被験者はルーレットを回してギャンブルを繰り返し行うギャンブル課題（左図）を行った。最初にどのような確率でどのような結果が得られるかを示す円形ボードが示され、その中心を矢印が回転する。一定時間後矢印が止まることで結果が決まり、ハイライトされた。したがって、予期効用と経験効用とを区別してしらべることができた。また、円形ボードに示される結果パターンには、損失と損得無しからなる“悪”条件、獲得、損失、損得無しからなる“中間”条件、そして獲得と損得無しからなる“良”条件の３条件があった（上図）。したがって、悪い結果の予期と良い結果の予期、あるいは悪い結果の経験と良い結果の経験とを区別することができた。前頭眼窩野、前脳基底部、線条体の一部（側坐核）が、予期効用と経験の両方を表現し、結果が良いほど、高い活動を示した。このことは、前頭眼窩野から線条体腹側部に到る領域が、予期効用と経験効用の両方を表現していることを示唆した。



「人はパンのみにて生きるにあらず」という聖書の言葉を引くまでもなく、私たちはジュースや食べ物やお金だけを求めているわけではありません。例えば他人から褒めて貰えるかどうかということも、私たちの行動を大きく左右します。このような他人からの賞賛が、食べ物やお金と同じような価値あるものとして、脳内で表現されていることも最近になって明らかになりました。Izuma ら{Izuma, 2008 #3795}は、自分や他人の顔写真と一緒に、人を評価する形容詞を呈示し、その望ましさを評価する単語評価課題（右図）を被験者に行って貰っている間に、fMRI を用いて脳活動を測定しました。すると、望ましい

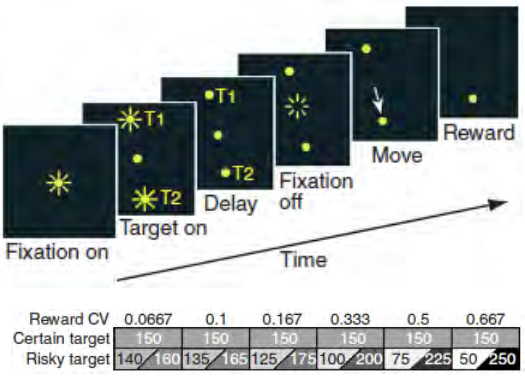


形容詞が自分の顔写真と一緒に呈示されたときには、お金を多く貰ったときと同じ線条体の一部が活動を高めた。

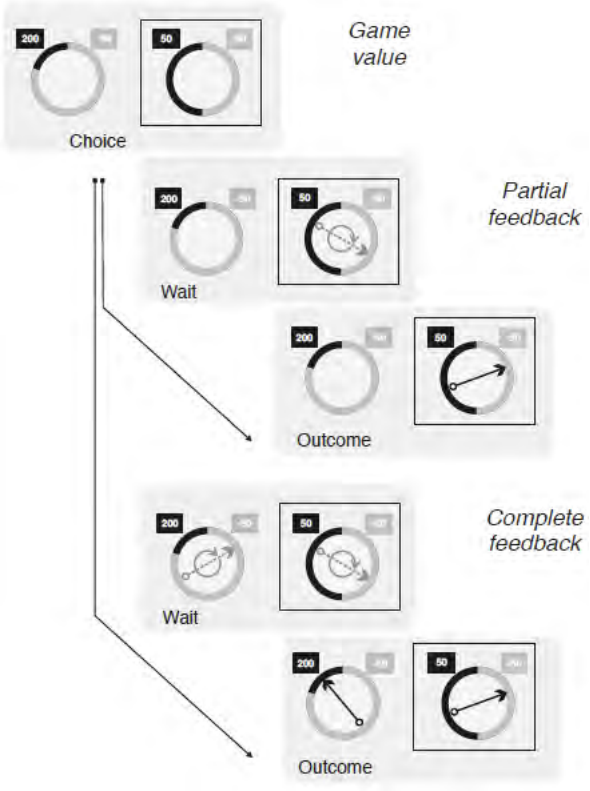
③リスク

人は一般に、不確実な状況での選択を嫌う傾向がある。この傾向も、従来の規範的経済理論では説明できない非合理的な側面である。

不確実さは、行動経済学的には、２種類に区別される。一つめは、それぞれの事象の起こる確率が、分かっている場合で、これを「リスク」と呼ぶ。それに対して、それぞれの事象の起こる確率が分かっていない場合、これを「曖昧さ」と呼ぶ。確率が分かっていなければ、チャンスレベルの確率を想定すると思われるが、実はそうではなく、「曖昧」な条件よりも、チャンスレベルの確率が分かっている条件の方を人は好む傾向がある。Hsu ら{Hsu, 2005 #3304}は、この「リスク」と「曖昧さ」とが、脳の中で実際に区別されて処理されていることを明らかにした。彼らは、選択結果に不確実さ（リスクの場合と曖昧な場合のいずれか）を伴うけれども多額の報酬を貰える選択肢と、確実だけれども少額の報酬しか貰えない選択肢の間での選択課題（上図）を被験者に行って貰い、課題遂行中の脳活動を、fMRI を用いて測定した。その結果、前頭眼窩野と扁桃体の活動が、曖昧さの度合いと正の相関を示し、背側線条体の活動が、負の相関を示した。しかも、前頭眼窩野に損傷のある被験者の選択行動には、曖昧さとリスクの両方に対する感受性が見られなかった。このことは、前頭眼窩野と扁桃体が、結果の曖昧な選択肢を避ける傾向を生み出すことに関わっていることを示唆している。



McCoy と Platt {McCoy, 2005 #3191}は、二点のターゲットの間で、得られる報酬の期待値は同じであるけれども、



一方は常に一定量の報酬が得られ、他方は多量の報酬と少量の報酬が５０％ずつの割合で与えられる条件で、サルにターゲットを選ばせる実験（左図）を行った。サルは、報酬が多かったり少なかったりという不確実性（確率が５０％と分かっていると思われるのでリスク）の高いターゲットをより高い頻度で選ぶ、つまりよりギャンブル好きの傾向を示した。この課題遂行中に、帯状回後部より単一神経細胞活動を記録したところ、それらの活動の大きさは、リスクの大きさ、つまり報酬量の分散（より正確には分散／平均）と相関した。このことは、帯状皮質後部が、リスクによる行動選好の修飾に重要な役割を担っていることを示唆している。

④後悔

人は、選択を誤って損害を被ったとき、がっかりしたり後悔したりする。自分の選択が思いもよらず悪い結果を引き起こしたときに私たちは「がっかり」し、もし別の選択をしていたならば、もっとよい結果を得られたことが分かったときに私たちは後悔する。Camille ら{Camille, 2004 #3800}は、この「がっかり」と「後悔」を区別した実験により、前頭眼窩野が「後悔」の中枢であることを明らかにした。彼らは、二つのルーレットのうち片方を選ぶ課題（上図）を 23 人の被験者にやってもらった。うち５人は前頭眼窩野に損傷のある患者さんで、残りの 18 人は健常な被験者であった。選ぶことによって回り始めたルーレットがどこで停止するかによって、お金を取得したり、失ったりする。取得および損失の確率は、ルーレットの外周に示された取得領域と損失領域の大きさによって示され、それぞれの領域には、取得額と損失額が示された。課題の条件としては二つ。一つは、選んだルーレットのみが回る条件。このとき、損失にルーレットが停止すると、健常な被験者は「がっかり」する。もう一つの条件は、選んだルーレットと選ばなかったルーレットの両方が回る条件。このとき、選んだルーレットは「損失」に停止し、選ばなかったルーレットは「取得」に停止すると、健常な被験者は、別のルーレットを選んでおけば良かったと「後悔」し、より、残念な気持ちになる。しかし、前頭眼窩野に損傷のある被験者は、二つの条件の間で、その残念な気持ちの度合いには有意な違いがなかったのである。このことは、選ばなかった方のルーレットの結果、つまりもしそちらを選んでおけば得られた結果についての評価に、前頭連合野が重要な役割を果たしていることを示唆している。

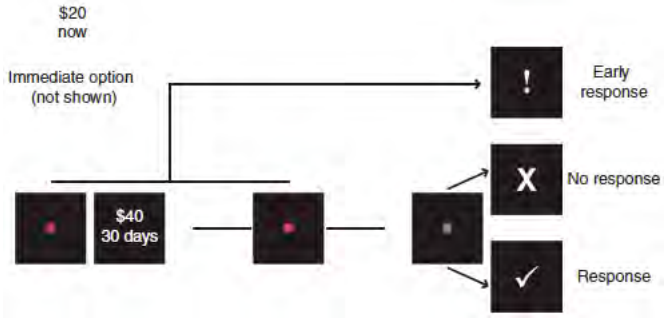
⑤時間割引

報酬が先延ばしされるとその価値は割り引かれる。例えば、直後の 1000 円と１年後の 1100 円のどちらを選ぶか尋ねると、大抵の人は、額が少なくても直後の 1000 円を選ぶ。この価値割引率がどのように決まっているかについては議論が続いているが、主に二つの考え方がある。一つは、指数型割引、もう一つは双曲型割引。経

済理論では、その簡便さから指数型割引の方が好まれてきたが、実験的には、むしろ双曲型割引の方に分がある。例えば直後の 1000 円と１年後の 1100 円との間で選んで貰うと上記のように直後の 1000 円を選ぶが、どちらも１年先送りして、１年後の 1000 円と２年後の 1100 円との間で選んで貰うと、逆に２年後の 1100 円を選ぶようになるという選択の逆転（選好逆転）が、起きる。このことは、割引率が一定となる指数型割引では説明できないが、時間が延びると共に割引率が減少する双曲型割引ではこれが説明できる。

しかし、直後の報酬の価値は、先送りされた報酬の価値とは、別々のシステムが関与しており、両者はそれぞれ別の指数型割引に従うという考えもある。McClure ら{McClure, 2004 #3029}の実験結果は、この考えを支持した。彼らは、本日、一週間後、そして一ヶ月後の報酬金の間での選択課題（左図）を被験者にやってもらい、課題遂行中の脳活動を計測した。本日（＝直後）の報酬金が選択肢として呈示されたときには、そうでないときと比べて、大脳辺縁系に含まれるいくつかの脳領域が、有意に高い脳活動を示した。一方、呈示された選択肢に含まれた先送り時間の長さに比例して、前頭前野外側部と頭頂連合野の各部位が活動を高めた。このことは、直後の報酬金の選択には、大脳辺縁系の各部位の活動が関与し、先送り時間が長い報酬金の選択には、前頭前野外側部や頭頂連合野の活動が関与し、両者の選択は別々の脳内システムに依存していることを示唆している。

Kable と Glimcher {Kable, 2007 #3760}は、先送りした報酬金の先送り時間を、より詳細に 1 日、10 日、21 日、50 日、90 日、180 日に分けて、直後の報酬金との間での選択課題（下図）を被験者に行って貰い、報酬金の主観的価値と脳活動との関係を調べた。その結果、腹側線条体、前頭前野内側部、帯状皮質後部の３つの領域で、主観的価値と相関した活動を示した。また、それらの脳領域の活動は、個々の被験者間での主観的価値の違いとも一致していた。



⑥損失忌避

上述のように、同じ額の利得と損失の可能性が５０％ずつのとき、その賭けに人は乗らない「損失忌避」の傾向が人にはある。Tom ら{Tom, 2007 #3686}はこの損失