

正負の個人的価値の理論としての感情価反応依存説 (VRD)

鈴木 真

私たちの多くは、次のことにたやすく同意する—実際それを主張する。私たちは禍 ill-being の理論 (あるいは、より十全な禍の理論) を必要としており、それは内在的な悪の存在を認め、それらの悪とは何なのか、そしてそれらがどれだけの重みをもつのか告げる (すなわち、その様々な悪がどれほど、そしてどんな仕方であらう者の福利水準を低下させるのかを示す) ものなのだ。それなのに、福利 well-being の典型的な議論においては、禍は大概無視されているように私にはみえる。それはほとんど全く注目をされていない。(Kagan 2014, p. 263)

1 はじめに—負の価値、絶対的な尺度、個人的価値—

直観的に考えると、価値には正の価値と負の価値があつて、それらは対称的なものである。対称的であるというのは、正の価値があればあるほどよいことであると同様、負の価値があればあるほどわるいことであり、そしてこのよさとわるさは一次元の尺度で表すことができ、よさの度合いは正の数で、わるさの度合いは負の数で、それぞれ表され、物事の総合的な価値はその剰余になるということである。しかしこの対称性の主張を正当化するのは必ずしも易しくないし、思想史的にも普遍的な支持を受けてきたわけではない。たとえば、悪というのは本当は善が欠如した状態に過ぎない、という立場をとる思想家は中世以来多くいたが、こうした思想では悪の度合いは負の数では表されず、低い正の数で表される。つまり、積極的な悪というものはないことになる。また負の価値を導入することによる理論的な複雑さが忌避されているのか、決定理論においても負の価値・効用があるという想定は—絶対的な0(ゼロ)があるという想定すら—普通はなされていない。

もう少し厳密に述べておくと、どのようなまともな価値理論においても、よりよい、という関係と同様、よりわるい、という関係も意味を成す。したがって、まともな価値の理論においては、無実の人が殺されるのは、そうでないよりもわるい、というこ

とが言える。このような相対的な悪の主張には特別な困難はない。このような主張では積極的な悪を認める必要は無いからだ。問題は、絶対的な負の値で表されるような価値があるという主張にある。たとえば、1回拷問されるより2回拷問される方がわるいというだけでなく、拷問されることは絶対的な意味でわるいことであり、たとえ殺されるか拷問されるかしか選択肢がない場合でも、拷問されるのは(死ぬよりはましであったとしても)わるいことには変わりがない、と私たちは考える。1億円か100万円をもらえるというくじがあって、100万の方のくじを選んだとしても、それがわるいとは考えない—1億円をもらえるよりは100万円をもらえることはわるいけれども、100万円をもらうことはわるいことだとは思わない—のとは明らかに違う。こうした絶対的な価値の違いについての直観を支持できるのかどうかというのが問題になる。

以下では、個人的価値 *personal value* の領域において、積極的な負の価値の居場所を確保し対称性の主張を正当化できるような負の価値の理論の方向性を提示しようと試みる。ここでいう個人的価値とは、ある主体¹⁾にとって—ある重要な意味でその者の視点から—価値あるもの、大事なもののことである。たとえば、

私が拷問されることは私にとってわるい

は、私という主体に相対的な価値の主張であるが、これは個人的価値の主張の一例である。これに対して、

無実の人が拷問される世界はそうでない世界よりもわるい

という主張は、「非個人的価値 *impersonal value*」ないし「世界にとっての価値 *value for the world*」と呼ばれる、主体に相対的でない価値についての主張の一例である。本稿では非個人的価値については議論しない²⁾。

個人的価値の理論は、福利の理論—ある者の生をよりよく、あるいはよりわるくする物事についての理論—を(少なくともその一部として)含む。個人的な価値の理論も福利の理論も、通常はある事態に内在する性質ないし関係自体のために価値があるということと、そのような事態が成り立つのを促したり妨げたりするために派生的に

価値があるということを区別して、前者の内在的価値の理論を提示しようとする（後者の手段的価値の理論はそれに基づいて得られる）ものであると考えられており、本稿もこの考えに従う。

以下では、まず福利の理論について、快樂説以外の説は対称性の説明に大きな問題を抱えることを示す。次に、快樂説の抱える難点を踏まえて、これを回避しつつ対称性の説明を提供できるような理論の方向性を提示する。そして実は、この理論は福利に限った価値の理論ではなく、より広く個人的価値の理論にもなっていることを指摘したうえで、そのような理論として擁護する。この理論は心理状態の感情価 *valence* に訴えるものになるので、感情価についてもある程度の考察をする。

2 代表的な福利の理論と負の福利の問題

福利の理論において負の価値—福に対するところの、禍—があるということは直観的にはもっともらしい。冒頭のケーガンからの引用にもあるように、福利の哲学的な議論において負の福利の扱いはあまり議論されてこなかったが、対称性自体はもっともらしいと考えられている⁽³⁾。しかしこのテーゼを正当化することは必ずしも楽ではない。福利の代表的な理論のいくつかについて、そのことを見ていこう。

福利の（客観的）リスト説は、福利の構成要素として複数のものを列挙する。このリストでは禍の構成要素が挙げられないことも多いのだが、挙げられる場合でもそれに統一性はない。たとえば、有名なパーフィットのリストを見てみよう⁽⁴⁾。

よいものの中には、道徳的善良さ、理性的活動、自分の能力の発展、子供を持ちよい親であること、知識、真の美しさの意識が含まれる。わるいものの中には、裏切られること、操作されること、中傷されること、騙されること、自由や尊厳を奪われること、サディスティックな喜びを感じることに、実際には醜いものに美的な喜びを感じる事が含まれる（Parfit 1984, p. 499）。

このように禍の構成要素は様々であり、それに共通する、それがなぜ当事者にとってわるいのか、ということを説明する性質はない、というのがリスト説の立場である。それらの複数の禍の構成要素—たとえば、自由や尊厳を奪われる度合いとサディステ

イックな喜び—が（当事者にとっての）わるさにおいて一元的な尺度で比較可能である、という保証はない。さらにいうと、様々な禍の構成要素と福の構成要素が一元的な尺度で比較可能であって、その尺度上で禍は負の値をとり福は正の値をとる、という保証もない。リスト説において対称性の主張がなされるとしたら、それは説明や正当化ができない事実として根拠なしに想定されるだけに過ぎない。

完成説 perfectionism（発達説 developmentalism）は、リスト説と違って、福利の統一的な説明を提供しようとする。この説によると、ある人の福利—あるいはその正の側面—は、その人の特異な、ないしは本質的な、機能を完成ないし発達させること、そしてそれを十全に発揮することに存する。しかしこの説明では、禍（負の福利）がどのように位置づけられるのかは明らかではない（Bradford 2016, p. 132）。一定の機能の発達や発揮の欠如が悪だ、と考えると、それは悪は単に善の欠如にすぎないという説となり、対称性テーゼを維持できない。もちろん、そのような発達や発揮を妨げる物事を禍（悪）と呼ぶことはできるが、これは単に手段としての悪であって、内在的にわるい事態を特定しているわけではない。

欲求充足説は、基本的に、ある者の福は欲求を実現する事態にあるとする。わかりやすく言えば、自分が求めているものが得られれば得られるほど自分の人生はよくなるという理論である。欲求充足説は 20 世紀以来決定理論・経済学において福利ないし効用の理論として主流の位置を占めており、哲学においても支持者が多い。決定理論や経済学における欲求充足説の理論では、（絶対的な意味で）負の価値・効用があるという想定は普通はなされていない。また最近までは哲学においても欲求充足説の下で禍の理論が語られることはまれであった。このような伝統的な形態では、福と禍の対称性のテーゼは主張されていない。

最近の哲学における欲求充足説の擁護者は、欲求が実現しない事態を禍とみなすというアイデアを採用することが多い。欲求の非実現ということをどのように特徴づけるかという難しい問題をわきに置くとしても⁶⁾、本当に欲求の非実現が積極的にわるいことなのかという根本的な問題がある。欲求が実現することが善であるとする、それが実現しないことは善の不在に過ぎず、積極的な悪とは言えないのではないだろうか（Kagan 2014, pp. 268）。また以下の点を考えてみてほしい。欲求充足説によると、（1）ある事態 X に欲求をもち、さらに（2）X が成立することが福なのだが、（2）の条件だけが成立しても禍が生じるわけではない。たとえば、（1'）私が今日の昼に

リンゴを食べたいと欲求して、(2') 実際食べられたら、私にとってよいことだが、そういう欲求を持たずにリンゴを食べても一方の条件である (2') しか成立しなくても—それは私にとって積極的にわるいことではない。単に私にとってよい事態が生じないだけである。だとすれば、今日の昼にリンゴを食べたいと欲求したが食べられない場合も一方の条件である (1') しか成立しなくても—それは私にとってわるいことではないのではないか (Kagan 2014, pp. 268-269)。

欲求充足説には禍を説明する別の可能性がないわけではない。たとえば、嫌忌 aversion という欲求と反対の「向き」を持つ心理状態を導入し、嫌忌される事態の成立を禍とみなすことも理論的には考えられる。この選択肢の問題については後述する。

3 快樂説の問題

代表的な福利の理論のうちでは、快樂説—快樂だけが福であり、苦痛だけが禍である—が絶対的な負の価値の居場所を見つけ対称性を統一的に説明するのに最も有望な理論である⁽⁶⁾。もちろん、快樂説でも快樂という正の価値と苦痛という負の価値がどのように比較評価の尺度に乗るのかということについての説明には問題が残る。快樂の度合 1 と苦痛の度合 1 が福利の価値尺度の上で絶対的に—その他の事情(たとえば、当事者が被っている快樂や苦痛の総量など)に依存せずに—相殺し合うものである、という Bentham (1781, Ch. 4) の立場は単純で魅力的ではあるが、その正当化は必ずしも簡単ではない。しかし、他の説はさらにひどい状況なので、負の福利の問題だけに注目するなら、快樂説には相対的な利点がある。

しかしよく知られている通り、快樂説は、私たちの多くが持つ客観的な状態や知識(あるいは一定の事柄に関する無知や誤解の欠如)に対する価値づけと両立できないという問題を抱える。人々が自分の福利に関係があると認める事柄は快苦以外にも数多く、しかもそれらは快苦を得る手段としてだけ重要視されているわけではないようにみえる。ノージックの有名な思考実験を振り返ってみよう。

あなたが望むどんな経験でも与えてくれるような、経験機械があると仮定してみよう。超詐欺師の神経心理学者たちがあなたの脳を刺激して、偉大な小説を書いている、友人をつくっている、興味深い本を読んでいるなどとあなたが考えたり

感じたりするようにさせることができるとしよう。その間中ずっとあなたは、脳に電極を取り付けられたまま、タンクで漂っている。あなたの人生の様々な経験を予めプログラムした上で、あなたはこの機械に一生繋がれているだろうか。

(Nozick 1974, p. 42)

この機械は経験をどのように変えることもできる。また機械に繋がれた後は、あなたは繋がれているという記憶をなくし、現実の経験をしていると思いつむようになるとする。経験機械が壊れたり不具合が起こったりして起こされることがないとすると、経験（たとえば、快楽と苦痛）にしか究極的には価値がないと考える説では経験機械につながれる方がよいことになる。しかし全てではないにしろ多くの人は、この選択が自分にとってよいとは考えない。彼らは、現実の正しい認識や、物事を実際に達成することや、人間関係を本当に築くことに福を（あるいはこれらが成り立たないことに禍を）見出しているようにみえる。

このような議論のために、福利について考える論者の多くは、快苦を福利の一要素とは認めても、それがすべてであるという快楽説には同意しない。私もこの理由には納得する—少なくとも、人によっては快苦以外の事態にも個人的価値がある—と考える—ので、快楽説をとらずに絶対的な負の価値の居場所を見つけ対称性を正当化したい。どのようにしたらよいだろうか。

快楽説が根本的に間違っていると思うのは、ある存在の快苦はその存在にとって重要なことのラベルであるにもかかわらず、そのラベルの方にだけ価値を認めているところである。進化論的に言えば、快楽や苦痛を感じる能力は、それが生存や繁殖に対して重要な事柄に注目を向ける働きがあるから私たちのうちに残ってきたのだと考えられる (cf. 廣中 2003, *passim*, e.g. 223)。実際、快苦（の情動的側面）にほぼ対応するものとしてのアフェクト *affect* に関わる脳回路は、恒常性や代謝に重要な情報に注目を向ける機能を持つと言われている (Clark-Polner, Wager, Satpute, and Barrett 2016, p. 152)。直観的にも、快苦だけに価値があると考えるのはおかしいようにみえる。たとえば私が腕を骨折したとすると、これは非常に痛い。ここで痛みを完全に鎮める薬があったとしよう。痛みをとるためには骨折を直さなくても薬を飲みさえすればよいが、それでいいと思う人がいるだろうか。いないだろう。骨折の方が痛みよりも本質的な問題だからだ。この例自体は快楽説の反例にはならないかもしれない。なぜなら

快樂説は骨折を直さないでいることによって不便が生じ、快樂が減ったり苦痛が増えたりする見込みを指摘できるからである。しかしもっと根本的に、無痛症の人がいるとしよう。この人は痛みを感じられないが、腕の骨を折ったら治療しないだろうか。その人の快樂の量に変化がないように医学的・福祉的な介入で調整できるとしても、それでも腕の骨を直した方がよいとほとんどみなが考えるだろう。こう考えていくと、快樂は当事者にとってよいことがあるということ、苦痛は当事者にとってわるいことがあるということを示しているという快樂説の着想は正しくても、本来的なよさやわるさはそうした状態が指し示している事態—ここでいうと腕の骨折—の方に(も)あると考える方が自然であるように思われる。この快苦はラベルであるという考えを用いて、客観的な事態を当人にとっての善や悪と認める理論を構築できないだろうか。

4 快樂説の代案

私の答えは、できる、というものであるが、その話をする前に快樂や苦痛というもののどのような側面が福利と特に関係があるのか考えなおしてみる必要がある。特に身体的な苦痛について明らかなことであるが、それには感覺的な側面と情動的側面がある。たとえば、あなたが風邪をひいて喉が痛いとする。この苦痛という心理状態は、のどというロケーションに炎症があることを示しているといえる。この苦痛はこの情報を意識にもたらししてくれるので、あなたはそれに気づいてのど飴をなめたり病院に行ったりといった対応をとることができる。この点で身体的苦痛は視覚や聴覚が外の環境の情報を運ぶのに対して、体の内部の情報を運んでいる内的感覺として働く。しかし身体的苦痛はそうした情報を運ぶだけではなく、独特な感じを伴う。このつらい感じは視覚などのただの感覺には必ずしも伴わないものである。普通の身体的苦痛の場合、この感じの側面と感覺の側面は離れることはないように見えるが、一定の脳障害を被った人は苦痛はあるがつかないという趣旨の発言をするという⁷⁾。この二つの側面は分離可能なものである。

さてそうすると、快樂説の論者や私のような快樂説から学ぼうとする価値論者にとって注目すべき側面は二つのうちどちらであろうか。快樂説論者にとっては、明らかに感じの側面であるように思われる。感覺的な側面は身体的苦痛によってもさまざまであり、さらに「精神的苦痛」というように呼ばれる、快樂説の論者が苦痛の範疇に

入れてきたもの、あるいは快樂には、必ずしも感覺的側面は伴わないかもしれない。しかしこれらは明らかに価値論の上からは少なくとも同様に重要であるようにみえる。

心理状態にのみ価値を認める快樂説にとっては、これで話は終わりかもしれない。しかし、快樂説でない説を求める私のような論者にとっては、感覺的側面も重要である。身体的苦痛は、身体のある個所の炎症を表すとともに、それについて感じを加えることで、それを評価に供するという役割を意思決定や状況の意味付けのうちで果たしているようにみえる。正確に言うと、苦痛が感覺的側面を持つこと自体が重要なのではなく、それが何かを表しているという側面—志向性—が重要なのである。快樂説論者が歴史的に「苦痛」や「快樂」と呼んできたものは必ずしも感覺的側面を持っておらず、より認知的な表象とも結びついていた。だから、「苦痛」や「快樂」には命題的態度—命題を志向の対象とする心理状態—として扱われるのが適切なものが含まれている。悲嘆の感情、たとえば人との別離に対する悲嘆は、「苦痛」として扱われることが多かったが、感覺的な側面を含んでいるようには必ずしもみえない—悲嘆には「胸が張り裂けそう」といった表現がなされることがあるので、もしかしたら、なんらかの身体の状態の感覚を含んでいるのかもしれないが。しかし、何らかの対象（この場合では、人との別離）の表象であるという点は身体的苦痛と変わらない。ここに「苦痛」の感じが加わることで、人との別れが個人による意思決定や状況の意味付けにおいてネガティブなものとして扱われるようになってきていると思われる。

「苦痛」や「快樂」の志向性の側面と区別される感じの側面は、心理学では感情価として理論化されている。感情価にはネガティブとポジティブの二つがある（どちらも持たない、ニュートラルというのを認めるなら三つである）^⑧。ネガティブな価をもつものが「苦痛」、ポジティブな価をもつものが「快樂」と考えることもできる。しかしこの定義は私たちの「快樂」や「苦痛」の概念を拡張するもののように見える—たとえば、欲求は（欲求されている状態に対する）ポジティブな感情価をもつようにみえるが、欲求自体を「快樂」と呼ぶことは変な気がする。また、（たとえば汚物に対する）嫌忌はネガティブな感情価を持つように見えるが、だからと言って嫌忌自体が「苦痛」だということも変である。思うに、普通の私たちの概念では、欲求や嫌忌や不安などの想像ないし予測された状態に対して持たれる心理状態は（予測やそうした想像をもつという事態に対する反応としての）快樂や苦痛を伴うことはあっても

それ自体としては快樂や苦痛ではない。快樂や苦痛というのは、何らかの起こっている事態に対する同時的ないしは事後的な状態—反応—だと考えるほうがもっともらしい。

この主張ですら苦痛や快樂の概念を拡張するものだろうが、本稿にとって重要なことは、この志向的な反応の状態と感情価という組み合わせが、事態が個人的価値を構成するかどうかの基準を提供するのではないか、ということである。すなわち、ポジティブな感情価をもつ反応の志向的对象となるものが正の価値を持ち、ネガティブな感情価をもつ反応の志向的对象となるものが負の価値を持つというように考えるのである（このアイデアには洗練が必要だが、今は単純な形で考えるほうがわかりやすいのでそうする）。こうすると、心理状態以外の事態に対して正と負の個人的価値を帰属することができる。こうした感情価を持つ反応が対象の評価を当事者の意思決定に供するものだとすれば、この個人的価値—当事者の視点からの善と悪—という理解はある程度納得できる面があると思う。強くポジティブな感情価の反応の対象になる事態はその反応を持つ人にとってそれだけ大きな正の価値をその時点において持ち、強くネガティブな感情価の反応の対象となる事態はその反応を持つ人にとってそれだけ大きな負の価値をその時点において持つ、とひとまず理論化することができる。

なお、ここでいう「反応」には刺激に対する反射のようなものだという含みはない。ある事態を受けて起こるアフェクトや情動は、大部分従来の経験による刺激の予想に基づくものであって、感覚刺激自体はその予想が違っていたら改訂する作用を限定的に及ぼすだけなのかもしれない（Barrett 2017, esp. Ch. 4）。しかしそうしたアフェクトや情動はそれでも感覚刺激によって調整を受けるものという意味で、純粋な予測や想像ではなく、反応であるといつてよいはずである。

ここでなぜある事態の純粋な予測や想像に対して持たれる心理状態を除外しているのかというと、それは一つには、その予測や想像が必然的に不精確であり、現実になり立つ事態の正確な評価を表しているとは思えないことが多いからである。たとえば、あなたは自分が付き合っている人と別れることを想像しても笑い話にして話だけでそれほどネガティブな態度を向けないかもしれないが、現実はその人と別れる事態に立ち至れば強いネガティブな感情を抱くかもしれない。どちらが成り立った事態—ある恋人との別れ—に対する評価として適切かといえば、それは別れた際の、あるいはその事後の、反応の方であろう。もちろん、予測ないし想像した事態に対する感

情が、その事態が実際に成り立った時の感情的反応をうまく「先取」していることも多いわけで、前者は後者の予測に使えるし実際私たちはよく使っている。しかし前者を後者の代わりにはできない。現実の事態を正確に想像・予測することも、その時に自分がどう反応するかシミュレートすることも、いずれも難しいからである。ここである事態の予測や想像に対して持たれる心理状態を除外しているもう一つの理由は、このような事前の予想に基づく心理状態—欲求、不安、恐れなど—は、動機づけに直接的に関わっており、当の本人からも他の人々からも操作されがちだからである。たとえば、あなたのパートナーと付き合いたい人は、あなたがパートナーと別れることを促したくて、その別れをバラ色に見せる—よい予測的感情を持たせる—よう、心理的操作を試みようとするかもしれない。あるいはあなた自身が仕事や勉強に集中したいからと、別れをバラ色に見せようと試みるかもしれない。こうなると、あなたのパートナーとの別れへの予想的感情はさらに実際の反応からはかけ離れていく可能性がある。

上記のコメントは、欲求充足説の支持者が嫌忌 (aversion) という欲求と反対の「向き」を持つ状態を導入することで、負の個人的価値を説明しようという試みをするのに対して、批判を投げかけるものである。欲求も嫌忌もある事態の予測や想像に基づいて持たれる心理状態であり、その事態の適切な評価を表しているとは考えにくい⁹⁾。嫌忌という心理状態が欲求と独立にあるという説を受け入れるとしても、そのような予想的心理状態は事態の価値の評価基準としてはふさわしくないとと思われる¹⁰⁾。

5 成り立ったことが認識されない事態の価値と、認識と注目の条件

上記の議論は、いかなる反応的心理状態でも価値基準となると考えるわけにはいかないことも示唆する。現実成り立った物事であっても、私たちはそれを正確に認識するとは限らない。不正確な認識に基づく反応はよくおこる。たとえば、夜遅くあなたが家の最寄り駅の改札から出ると一人の男が立っていたので怖がったとしよう。この男があなたのパートナーであなたが遅いから心配して改札で待っていたのだとすると、本当のところ正確に理解されたところのこの事態に対してネガティブな反応をするいわれはない(夜に誰かわからない人に直面するという状況だということについては、あなたは正確に認識しているので、それに対してはそうした反応が適切かもしれ

れないが)。このように言うためには、反応が価値の基準となるのは、状況を正確に把握した場合のみだという条件が必要になる。

ところで、成り立った事態に対する反応を基準とすると、単純に考えれば、認識されない事態には価値が帰属されないことになる。経験機械に入った人は、現実には自分が経験していると思っていることが成り立っていないということを認識しないし、経験機械に入っている間に健康がむしばまれたり家族に見捨てられたりしていても認識しない。そうすると、こうした事態に対しては（個人的）価値が帰属されないことになる。価値が帰属されるのは、経験機械のうちでの経験—たとえば、いいパートナーとキャリアと子孫に恵まれ充実した生を送っているようにみえるということ—だけであって、それ以外の現実には起こっていることには価値が帰属されない。そうすると、多分そうした経験はポジティブな反応の対象になるから、経験機械に入ることは望ましいことしかないという結論になってしまうだろう。これは快樂説を退けようとする企図に反する。前段落において述べたように、正確な認識が必要とされるなら、なおさら価値が帰属されるが限定されるので、問題が悪化する。

そうした認識されない、あるいは正確には認識されない、事態も価値を持つと論じるには、反事実条件文による条件設定に頼る必要があるだろう。すなわち、もし仮にそうした事態が正確に認識されたとしたら、当事者のポジティブな感情価を持つ心理状態の対象となるであろうものがその存在にとって正の価値を持つものであり、ネガティブな感情価を持つ心理状態の対象となるであろうものがその存在にとって負の価値をもつものである。現実に正確に認識される場合は、条件文の前件（条件節）が真となるので、後件が真かどうか経験的に調べることができる。この条件節が現実には真とならない場合には、当事者についての情報を集めて分析したり、シミュレーションをしたりして推測するしかない。

上のように考えると、経験機械に繋がれている生より現実の生の方がよいかもれないと考えられるわけが説明できる。経験機械に繋がれている人は実際には現実の事態の多くを認識しないわけであるが、もし仮に認識したとしたら、その人は（もし私と似ているなら）多くのことについてネガティブな感情価を持つ反応をするであろう。たとえば、あなたは自分が現実には何も達成しておらず、誰も友人がおらず、そのことを知りもせずいたということに対して、それぞれ悲しんだり悔しがったりといったネガティブな感情価を含む反応をするであろう。これらのことは、少なくとも、経験

機械に繋がれた人生と、それがヴァーチャルに経験させてくれることを実際に経験する人生とを比べたときに、後者の方がその人にとってはよりよい(よりわるくない)、ということの説明しうる。快樂説と違って、この説—感情価(をもつ)反応(への)依存説 Valenced-Response-Dependence Theory、略して VRD とでもよぶべきもの—は経験機械の思考実験に基づく批判をかかわることができる。

上で言及した認識上の条件をより厳密に述べるなら、事態の認識の正確性だけではなくそれへの注目も要求される。あなたのパートナーや友人があなたに忠実であることは正確に認識されていたとしても、普段あまり注目されないために何の反応も呼び起こさないかもしれない。だからといって彼らが忠実であることがあなたにとって何の個人的価値も持たないというのはおかしい。事態は正確に認識されるだけでなく、注目もされなければ、適切に反応されることはできないだろう⁽¹¹⁾。

ここで注意しなければならないのは、感情価をもつ反応は必ずしも注目が向けられた事態の個人的価値の基準ではないということである。私たちは時折反応の志向的对象となる事態ではなくて他の事態に注目してしまい、それが何に対する反応かを間違えて理解してしまうことがある。たとえば、人々は雨の日に自分の生の状態について問われると、晴れの日より低い満足度を表明する傾向があるが、この傾向は先に天気について明示的に問われるときには消える (Schwarz and Clore 1983)。これは人々が、本当は天気の状態に対するネガティブな反応であるものを、質問によって注目を向けていた自分の人生の状態に対する反応だとみなしてしまうことを示唆している⁽¹²⁾。ここで反応の志向的对象がどのように決まるのかについての理論を展開することはできないが (e.g. Prinz 2004, Ch. 3)、少なくともそれは人々の注目の対象と必ずしも同一ではない。感情価をもつ反応がその志向的对象となる事態の個人的価値の基準となるのは、他の事態ではなくその事態が正確に認識され、注目された場合であろう。

6 主観的経験自体の価値と高階の反応

VRD によると、感情価を持つ反応が個人的価値の基準となるのはその反応の志向的对象となる事態である。ここで、快樂説の論者でなくても、反応の経験的側面—たとえば、快樂や苦痛の感じの側面—はそれ自体として当事者にとって価値があるのではないか、と言いたくなるだろう。実際、緩和ケアのように身体的苦痛の志向的对象

とみなされる箇所が治療不能な場合でも苦痛だけをとりとうすることに意義があるとみなされるのだから、苦痛自体に負の個人的価値がないというのは信じがたいだろう。また人々が快楽を求めて様々な行動をとることを考えると、快楽自体にも正の個人的価値があると考えるのが自然である。

VRD はこうした直観を救うことができる。ある程度の強さを持つ苦痛や快楽は、それ自体として感情価を持つ反応の志向的対象となる。たとえば、私が癒えることのないガンによるひどい痛みを感じていれば、その痛みに対する負の反応を持つのが自然である。また、私があるリンゴを食べて強い快楽を感じたら、その快楽に対する正の反応を持つのが自然である。こうした高階の反応により、それらの苦痛や快楽の体験自体が（個人的）価値とみなされることになる。

もちろん上記のようにならない場合もあるかもしれない。ある人はポルノをみることや差別的なジョークを見聞きすることに快楽をおぼえるかもしれないが、その快楽の体験には負の反応をするかもしれない。そうすると、その人に関していえば、ポルノを見ることや差別的なジョークを見聞きすること自体は正の価値を持つが、そこから得られる快楽は負の価値を持つ。このように VRD では快楽説と違って、快楽は正の価値を持ち、苦痛は負の価値を持つ、というわけでは必ずしもない。しかしこれももっともらしくないかどうかは別問題であり、快苦についての個別主義的な直観を拾い上げるという点では VRD が勝っている。

高階の反応は常に快苦に伴うわけではないから、快苦に常に（正であれ負であれ）価値があるわけではないことになる、という批判もありうる。もちろん、すべての快楽や苦痛に高階の反応が現実に伴うわけではない。注目されないがゆえに高階の反応を呼びおこさない快苦もあるだろう。またすべての快苦に高階の反応が伴うと考えると、有限な私たちが無限に多くの心理状態を持つことになってしまうが、これももっともらしくない。しかし VRD では、ある事態が個人的価値を持つのに実際に反応が生じなければならないわけではない。VRD によると、もし仮にそうした事態が正確に認識され注目を向けられたとしたら、当事者のポジティブな（ネガティブな）感情価を持つ心理状態の対象となるであろうものがその存在にとって正の（負の）価値を持つ事態である。現実には反応されない快苦であっても、ほとんどのものは正確に認識されて注目されればポジティブないしネガティブな反応を伴うだろうと考えられる。あなたの頭痛は注目されないがゆえに反応を引き起こさないかもしれないが、そ

れが正確に認識され注目されればあなたのネガティブな感情価を持つ心理状態の対象になるだろうから、それはあなたの視点からはわるいものだと言えるのである。

なお、経験のうち快苦だけが感情価を持つ高階の反応の対象となるわけではないかもしれない。猫の肌触りや体臭自体は快樂ではなく、ほとんどの人にとっては何の反応ももたらさないかもしれないが、一部の人においては非常に強いポジティブな反応の対象となる。彼らにとってはそうした経験は正の個人的な価値があると言っていいだろう。

7 福利を超え出るものとしての個人的価値とその理論としてのVRD

VRD と似た構造を持つ理論は、福利の理論の外では何度も提示されてきた。Perry 1926 から Zimmerman 2011 に至るまで、ポジティブな反応とネガティブな反応を参照して価値の正負を説明しようとする哲学者は少なくない⁽¹³⁾。この方針を福利の理論において採用するのに特別な支障があるようには思われない。しかも、非個人的価値、世界にとっての価値についての主張についてこうした方針をとると直面する問題を考えると、この方向性は福利を含む個人的価値の理論の領域の方が展開しやすいようにみえる。

たとえば⁽¹⁴⁾、太郎と次郎が一つしかない大学のポストの公募に応募するとする。太郎と次郎には学問的に（も人柄にも経済的・社会的状態にも年齢にも）差がないとすると、太郎がポストを得ることも次郎がポストを得ることも等しくよい、あるいは、太郎がポストを得られないことも次郎がポストを得られないことも等しくよくない、という非個人的価値、世界にとっての価値の主張が正しいように見える。しかし、太郎や次郎がどのような反応をするか考えてみよう。太郎も次郎も自分がポストを得られたら喜ぶ—ポジティブな反応をもつ—だろうが、相手がポストを得たら悲しんだり悔しがったりする—ネガティブな反応をもつ—だろう。太郎や次郎の反応は上記の非個人的価値、世界にとっての価値の主張を裏書きするようには見えない。一方、この状況では、太郎にとっても次郎にとっても、自分がポストを得られることは自分にとってよいことだが、相手がポストを得ることは自分にとってわるいことである。この個人的価値の主張は、太郎や次郎の反応と一致するから、それらの反応によって裏書きされていると考えても不自然ではない。VRD のような理論は、個人的価値の理論と

してまずはもっともらしいようにみえる⁽¹⁵⁾。

しかし VRD が当事者の福利（利害）に限定された理論として提出されるのであれば、この説はあまりもっともらしくないのであろう。というのは、私たちは様々な事態に対して感情価を含む反応をし、それらは普通に考えると当事者の福利とは関係ないようにみえるものも含まれるからである。もちろん、私は自分がけがをしたりヒヤッとするような体験に遭ったりしたら、それに対してネガティブな感情価を持つ反応—前者では身体的苦痛、後者ではいわゆる「恐怖」や「怒り」—を持つ。これらの反応は私の福利の価値の基準となると考えてもよいと思う。しかし、私は美しい絵画や音楽や自然の景観の在り様に好意的な反応をし、貧しい人の経済状態が改善されたら喜び、誰か知らない人に対するオンライン上のバッシングに対して怒る。これらは私の福利と関係があるとはあまり思えず、たとえあるとしても反応の感情価の大きさが私にとっての福利の価値の大きさの基準だと考えるのは説得力がない。政府が私から税金をとって貧しい人に再分配するとき、その政策は貧しい人の経済状態を改善しているから私に利益を与えているのだ、というのはあまりもっともらしくない。また私はオンライン上のバッシングに対して非常に怒るかもしれない—非常にネガティブな感情価を持つ反応をするかもしれない—が、だからといってその見知らぬ人に対するバッシングが私の福利を大きく損なうと考えるのは変である。私個人の人生の成功は、見知らぬ人へのバッシングの有無にそれほど大きく依存していない。

VRD は福利に限定された価値理論ではなく、個人的価値一般についての理論であると考えるのがもっともらしい。（第 5 節で述べた条件の下では）ある主体に本質的に関わる事態⁽¹⁶⁾に対するポジティブないしネガティブの感情価を含む反応はその主体にとっての福利としての価値の基準だとみなすことができるが、その主体に直接関わらない事態についての反応は福利以外の個人的価値の基準だとみなすのが適当であろう。私にとっての個人的価値は、私の視点から見た価値であって、それには私の人生をよくいかせたりわるくいかせたりする事態も含まれるけれども、それ以外の事態も含まれるのである。私の視点から見たときの、美的価値、政治的価値、認識的価値などといったものも個人的価値には含まれる。VRD はこの様々な個人的価値を統合的に説明しようとするものであり、とりわけそうした価値の正負を感情価で説明しようとする理論なのである。

8 感情価

VRD のように心理反応のポジ・ネガによって価値の正負を説明しようとする理論は、感情価について一定の説明をする必要がある。心理学では情動はよく生理的な覚醒度 arousal とともに感情価 valence で区別される。つまり、感情価は情動の根本的特徴の一つだと考えられている。現代における代表的な情動の哲学者 Jesse Prinz(2004, Ch. 7)によると、この考えの根拠の一つは、情動の類似性空間の研究である。心理学者は、人々が情動についてどのように考えているかを知るために、実験参加者に情動語のペアの類似性を判断してもらい、統計学的手法を使って図式化している。より似ていると評価された情動ほど近くに配置され、円形構造が提示される (Larsen and Diener 1992; Russell 1980)。研究の積み重ねの結果、この円形構造は二つの対比次元に沿って組織化されるようになった (図1)。この次元が覚醒度と感情価である。なおこの図では情動カテゴリ/タイプが配置されているが、一つのタイプに属するトークンがすべて同じ覚醒度と感情価をもつという想定がなされているわけではない。たとえば同じ「恐怖」のトークンでも、興奮したものから穏やかなものまで、非常に不快なものから快いもの (スカイダイビングへの恐怖?) まで、あるかもしれない。

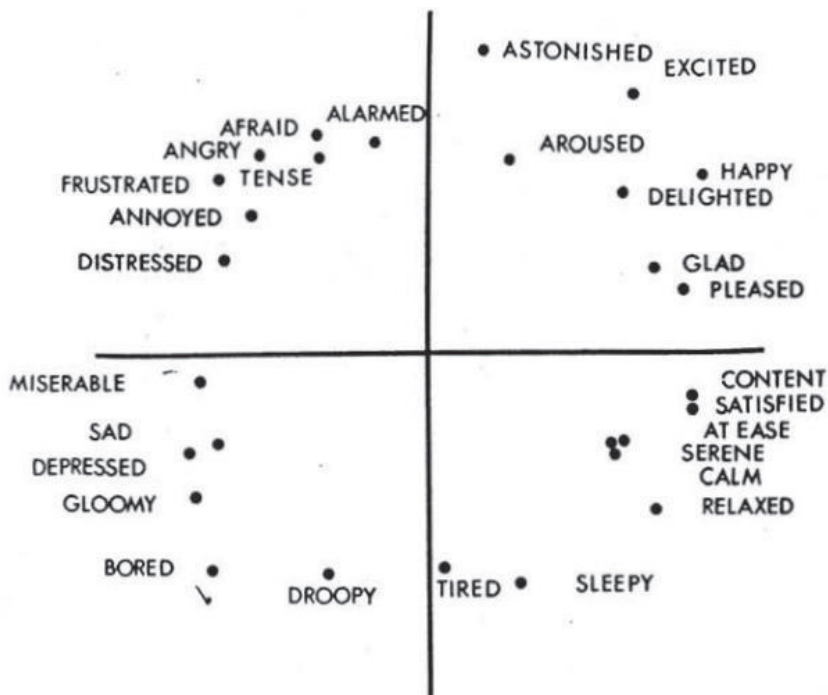


図1 情動/アフェクトの円環モデル (Russell 1980, Figure 4)

横軸が「不快 - 快」(感情価におけるネガティブ - ポジティブ)、縦軸が覚醒度を示す。

なお戸田山も指摘しているように (戸田山 2016, pp. 45-46)、この円形構造に提示されている状態には情動と呼ぶのが不適切にも思われるもの—「平静」・「落ち着き」のように気分ではないかとみえるもの、「眠さ」のように生理的反応のように見えるもの—が挙げられている。感情価を持つことは情動の必要条件であるかもしれないが (Prinz 2004, Ch. 7)、すべての感情価をもつ反応が情動であるわけではないだろう。

感情価という概念を価値の理論に持ち込む際には二つの懸念があるように思われる。一つは、感情価が普遍的に存在するものなのかという点である。この点を哲学者はこれまで真剣に考えてこなかったが、情動カテゴリ/タイプは実在しない—日常的な情動語「怒り」「悲しみ」「幸福感」「驚き」「嫌悪」などに対応するような、相互排他的に区別される心理的・脳科学的まとまり (自然種) は実在しない—という立場が

近年有力になってきていることや (e.g. Barrett 2006)、情動 emotions というカテゴリが他のカテゴリと区別されるまともりなのかどうかという点についても疑念が示されていること(Griffiths 2004)を踏まえると、無視できる問題ではない⁽¹⁷⁾。しかし感情価 (あるいはそれにおよそ対応する快樂や苦痛といった心的カテゴリ) については、少なくとも心理的な次元では普遍的に存在するという見解が有力である (Barrett 2006; Barrett 2016, p. 72f; Scarantino 2009)。これはたとえば、顔の筋肉の動き、経験の報告、末梢神経系の活動が感情価 (と覚醒度) と強い対応を示しており (e.g. Lang et al. 1993)、幼児にもこうした証拠がみられることによる (Mesman et al. 2012)。

ただし VRD は価値の程度を語るために感情価の程度について語ろうとすると、感情価に程度があるという更なる主張を擁護しなければならない。また、個人的善と個人的悪の対称性を擁護するには、正と負の感情価が同一尺度上で相反するものであることを擁護する必要がある。図1では前提されているこれらの点について、本稿では議論する用意がない。さらに脳科学では感情価がどのように実現されているのかについて論争があり、近年有力なモデル (affective workplace hypothesis) では、正負の感情価は状況によって様々な、重なり合う脳領域のまともりによって実現されているとされている (Clark-Polner, Wager, Satpute, Barrett 2016, pp. 147-152)。この仮説が正しいとしても、色々な心理状態が、正 (あるいは負) の感情価をもつことで、一様の(他の心理状態からは区別される)機能を果たすことは否定されないので、心理学的なレベルでは正 (と負) の感情価は一つのまともりとして実在しうる。しかし、もしかしたら感情価には実現のされ方によって機能の違いがあるかもしれない、この場合には様々なタイプの感情価があることになる。もしそうだとすれば、VRD を採用して事態を感情価をもつ反応によって評価しようとする場合、相異なるタイプの感情価をもつ反応の志向的対象となる事態をどのように統合的に評価するかという問題に直面することになるだろう。だが、この質的快樂説が直面するのと似た問題についても、これ以上論じることはできない。

もう一つの懸念は、何が正の反応を正のものに、何が負の反応を負のものにするのか、という感情価の本質にかかわる事柄である。残念ながら、この本質を解明する理論についてはコンセンサスがない。ここで本稿にとって懸念される材料は、もし感情価の正・不正というものが、価値の概念や価値判断に依存するものだとすると、感情価によって個人的価値を規定しようとする VRD は循環する恐れがあるということで

ある。そこまでいかなくても、VRD は価値を価値概念や判断に依存させるという意味で一種の客観性を放棄することになるだろう。幸いなことに、感情価を価値概念や価値判断に基づいて理解しようとする説は人気がある説ではない。これは、感情価を持つとされる心理状態が大人の人間だけでなく価値概念や価値判断を下す能力を持たないとされる幼い子供や多くの脊椎動物にも共有されている⁽¹⁸⁾というだけでなく、感情価を還元的に定義する様々な可能性があるからである。Prinz (2004, Ch. 7) は以下のような説を批判的に検討している。

- ポジティブな感情価は快樂に、ネガティブな感情価は苦痛に存する
- ポジティブな感情価は目標と合致することに、ネガティブな感情価は目的に反することに存する
- ポジティブな感情価は接近の傾向に、ネガティブな感情価は撤退の傾向に存する
- ポジティブな感情価は外への注目の促しに、ネガティブな感情価は内への注目の促しに存する

Prinz 自身 (2004, pp. 173-174) は、Jeffrey Gray (1991)の節を改良して、内的強化子を感情価と同一視する。彼によると、脳には二つの内的強化子—刺激の表象と結び付けられた心理状態—があり、これらが感情価マーカーなのである。内的な正の強化子は「これを増やせ!」、内的な負の強化子は「これを減らせ!」、というようなことをいう命令としてそれぞれ働く。「これ」の指示対象は「身体性の評価」表象⁽¹⁹⁾との組み合わせによって与えられる。感情価を持つ状態とそれに伴った「身体性の評価」表象とともに記憶に蓄えられると、意思決定に影響を与えることになる。

Prinz によると、内的評価子の正負は情動が表す価値的なあり方の正負とは独立である。信原 (2017, pp. 197-200) はこの点を批判して、感情価は情動が表す価値的なあり方の正負に由来すると論じている。この見解—感情価の「価値表象説」—が感情価の正・不正が価値の概念や価値判断に依存するなら、この見解が正しければ VRD は循環することになる。しかし感情価が価値の表象だとしても、それは価値の概念や価値判断を前提するものではないだろう。そう考えてしまうと、感情価を持つとされる心理状態が価値概念や価値判断を下す能力を持たないとされる幼い子供や多くの

脊椎動物にも共有されているという事実と整合しなくなってしまう。また私たちは、何かをわるい（よい）と判断しなくてもそれに対してネガティブ（ポジティブ）な反応をもつことがあるように見える。だから、感情価が価値の表象だとしても、それは価値の概念や価値判断を前提するものではなく、より原始的な表象であると考えるのがもっともらしいだろう。

感情価が何に存するかという問題は VRD にとって核心的に重要だが、本稿ではこの点についてこれ以上述べる紙幅がない。

9 おわりに—残された課題—

本稿では、感情価反応依存説（VRD）が個人的価値の理論、特にその正と負の価値の対称性を説明する理論として見込みがあることを示そうとしてきた。それは快楽説と同様に正の価値と負の価値のそれぞれを統一的に説明し一定の経験に価値を認めるが、快楽説のように福利（やそれを含む個人的価値）を狭くとらえることなく、経験機械のような思考実験における私たちの直観も説明できる。また VRD は、私たちの情動的反応が自分にとっての価値の特定と関係している、という多くの人々の直観（Scarantino 2016, pp. 31-32）を救う。

ただし、VRD による対称性テーゼの説明の試みは、快楽説による説明と同様、端緒にすぎただけである。快楽説の論者は、快楽という正の価値と苦痛という負の価値がどのように比較評価の同一尺度に乗るのか、ということの説明をしそこなってきたが、VRD も今の段階ではこの説明ができていない。VRD によって対称性テーゼが正当化できるか否かは、感情価の本性についての実証的研究とともに、このような説明がうまくいくかどうかによるであろう。

VRD をきちんと展開するためには、感情価をもつ状態の志向的対象となる事態がどのように決まるのかについての理論も検討しなければならない。どのような理論が正しいかによって、VRD によるとどのような事態が当事者にとって価値があることになるのかが決まることになるからだ。

また当然ながら、VRD は個人的価値の規範性に関する直観（Scarantino 2016, *ibid.*）も説明することができなければならない。この説明要求にこたえるには、個人的価値の規範性がどのようなものであるかを明らかにして、私が支持する型の VRD を本稿

でした以上に詳述する必要がある。さらに、個人的価値と、非個人的価値、そして合理性や倫理との関係がどうなるのかという点についても論じる必要がある。

これらの課題⁽²⁰⁾には別の機会に対処することにした。

本研究は JSPS 科研費 16K02148 の助成を受けたものです。

注

- (1) 紛らわしいが、これは必ずしも人間や行為者でなくてもよい。個人的価値について語ることが意味をなすのはどのような存在者についてなのか、ということ自体が一つの論争点である。本稿の立場については注 18 をみよ。
- (2) 個人的価値と非個人的価値の区別とそれに関する論争については、たとえば Ronnow-Rasmussen (2011) を参照。
- (3) たとえば Sumner (1996, p. 13) はこうした対称性を含意し説明することが適切な福利理論の条件の一部だと考えているように見える。
- (4) パーフィット自身は以下のリストを例示のために使っており、リスト説を支持しているわけではない。
- (5) 特に問題になるのは、欲求の対象となる事態やその実現が求められる時点が当人によって抽象的ないし曖昧にしか特定されていない場合である。たとえば、私がブルーストの『失われた時を求めて』をいつか読破したいという欲求を持ってきたが現在まで読めてこなかったとして、これは欲求が実現されない（ので私にとってわるい）状態だとみなすべきだろうか。それとも私が死ぬまでにこの本を読めなかったときに限り欲求が実現されない（ので私にとってわるい）状態だとみなすべきだろうか。私がこの長い本を途中までしか読めなかった場合はどう考えたらよいだろうか。
- (6) 他には Sumner 1996 らの生満足説 life satisfaction theory がある。生満足説に対する批判については、たとえば Feldman 2010, Ch. 5 を参照。
- (7) この苦痛の失象徴 pain asymbolia については、Grahek 2007 に詳しい議論がある。
- (8) 本稿に関わる限りでの感情価についての哲学的問題については、第 8 節で触れる。
- (9) これは、Sumner (1996, p. 129) が指摘している欲求の予期性 the prospectivity of desire の問題である。Suzuki 2011 も参照。
- (10) とはいえ、この区別は時に直観的には難しく、理論的に引かれる必要がある。たとえば私が（映画や小説などのフィクションの対象にではなく）現実の状況について恐怖や不安を持つとき、これはある未来の予想された事態（たとえば、

交通事故にあうこと)についての心理反応にも、その事態を引き起こしうる現状の危険ないし脅威(たとえば、私の住んでいる名古屋では交通事故死者が全国的に多いということが傾向として持っている危険ないし脅威)に対する反応にもみえる。この二つの区別は根本的にはその心理状態の志向的対象がなんであるかということについての理論に依存するだろう。すなわち、ある恐怖や不安が未来や架空の事態についての心理状態だとそうした理論によって判明するならば、それは予想的心理状態であるが、(私がよりもっともらしいと思うように)それらが現状の危険ないし脅威に対する反応であるとそうした理論によって判明するならば、それは現実に対する反応である。

- (11) Lewis 1989 の価値の理論のように、想像上でできる限り鮮明に想起することまでを求めるのは行き過ぎのように思われる。そのような苦労と困難を伴う手続きを踏むことを多く人は通常していないので、そのような条件下における反応は、人々の実際の反応とは大きく違いうるし、その対象となる事態が通常の状態で人々の心を動かすということも一般的には期待できない。ルイスが記述するような手続きは、ある人の個人的価値について発見する手段というより、本人の価値に関する反応やそれに伴う考え方を大きく変える可能性を持つ手段だと見た方がよいように思われる。
- (12) 様々な例については、Barrett 2016, pp. 74-77 を参照。
- (13) Tappolet (2015, 91)もこのような理論をその基本的な課題とともに示唆している。その課題は、何が正の反応を正のものに、何が負の反応を負のものにするのか、という感情価の本質にかかわる事柄と重なる。この話題については第8節を参照。
- (14) 以下の問題の記述は、Zimmerman 2011 が価値の「適合態度分析 fitting-attitude analysis」に対する「不公平性 partiality による異議」と呼ぶものを、そうした分析だけでなく価値を心理状態に依存させようとする試み一般に当てはまるように翻案したものである。
- (15) Perry 1926 は価値の一般理論を提示しようとしながら、価値には主体に相対的なものしかないという立場を擁護することになった。Zimmerman 2011 は、基礎的な内在的価値と基礎的でない内在的価値を区別することを通じて非個人的価値が反応に基づいて分析できると論じている。
- (16) 別の言い方をすると、ある主体への言及なしに記述できない事態。
- (17) これらの点を考えると、事態の価値を「情動」というカテゴリに依存させたり、(たとえば Toni Ronnow-Rasmussen (2011)が様々な個人的価値についてしているように)「怒り」や「恐れ」などの下位カテゴリに依存させたりするのは危ういように思われる。それらのカテゴリは、実在しないなら、事態の価値を説明することはできないだろう。VRD はこれらのカテゴリには依存していない。
- (18) VRD を採用するならば、感情価をもつ反応をする存在については福利を含む個人的価値を語ることが意味をなすと考えるのが自然なので、大人の間人だけでなく幼い子供や多くの脊椎動物にとってもよいこと、わるいことがあると考えることになるだろう。

- (19) 「身体性の評価 embodied appraisals」というのは、一連の神経メカニズムが身体状態（の変化）を知覚し、そうすることによって生物と環境の関係（の変化）を表象している、という心理状態のことである。
- (20) 他にも様々な問題がある。たとえば、私たちの反応の感情価がある程度対象の提示のされ方に依存することや、馴化や嗜癖・依存によって反応が変わることや、感情価を持つ反応の認知的・情緒的基盤が障害を被る可能性などを考慮すると、VRDをさらに洗練する必要があることは明らかである。

参考文献

- Barrett, Lisa Feldman (2017) *How Emotions Are Made: The Secret Life of the Brain*. Macmillan. (高橋洋訳 (2019) 『情動はこうしてつくられる—脳の隠れた働きと構成主義的情動理論』 紀伊国屋書店)
- Barrett, Lisa Feldman (2006) “Are Emotions Natural Kinds?” *Perspectives on Psychological Science* 1 (1): pp. 28–58.
- Bentham, Jeremy (1789) *An Introduction to the Principles of Morals and Legislation*.
- Bradford, Gwen (2016) “Perfectionism.” In *The Routledge Handbook of Philosophy of Well-Being*, ed. by Guy Fletcher, Routledge. pp. 124-134.
- Clark-Polner, Elizabeth, Tor D. Wager, Ajay B. Satpute, and Lisa Feldman Barrett (2016) “Neural Fingerprinting: Meta-Analysis, Variation, and the Search for Brain-Based Essences in the Science of Emotion.” In Lisa Feldman Barrett, Michael Lewis, and Jeannette M. Haviland-Jones (eds.) *Handbook of Emotions* Fourth Edition, The Guilford Press, pp. 146-165.
- Feldman, Fred (2010) *What Is This Thing Called Happiness?* Oxford University Press.
- Grahek, N. (2007) *Feeling Pain and Being in Pain*. Cambridge, Mass.: MIT Press.
- Griffiths, Paul E. (2004) “Is Emotion a Natural Kind?” In R. C. Solomon (ed.), *Thinking about Feeling: Contemporary Philosophers on Emotions*. Oxford University Press, pp. 233–249.
- Gray, Jeffrey A. (1991) “The Neuropsychology of Temperament.” In J. Strelau and A. Angleitner (eds.), *Explorations in Temperament*, Plenum Press, pp. 105-128.
- Kagan, Shelly (2014) “An Introduction to Ill-being.” *Oxford Studies in Normative Ethics*, Volume 4: pp. 261-288.

- Lang, P. J., M. K. Greenwald, M. M. Bradley, and A. O. Hamm. (1993) "Looking at Pictures: Affective, Facial, Visceral, and Behavioral Reactions." *Psychophysiology*, 30: pp. 261-273.
- Larsen, R. J. and E. Diener (1992) "Problems and Promises with the Circumplex Model of Emotion." *Review of Personality and Social Psychology*, 13: pp. 25-59.
- Lewis, David (1989) "Dispositional Theory of Value." *Philosophy and Phenomenological Research* 51: pp. 113-137.
- Mesman, Judi, Harriet Oster, and Linda Camras (2012) "Parental Sensitivity to Infant Distress: What Do Discrete Negative Emotions Have to Do with It?" *Attachment and Human Development* 14(4): pp. 337-348.
- Nozick, Robert (1974) *Anarchy, State, and Utopia*. Basic Books. (嶋津格訳 (1988) 『アナキー・国家・ユートピア—国家の正当性とその限界』 木鐸社)
- Parfit, Derek (1984) *Reasons and Persons*. Oxford: Clarendon Press. (森村進訳 (1998) 『理由と人格』 勁草書房)
- Perry, R. B. (1918) *General Theory of Value*. New York: Longmans, Green.
- Prinz, Jesse J. (2004) *Gut Reactions: A Perceptual Theory of Emotion*. Oxford University Press. (源河亨訳 (2016) 『はらわたが煮えくりかえる—情動の身体知覚説』 勁草書房)
- Ronnow-Rasmussen, Toni (2011) *Personal Value*. Oxford University Press.
- Russell, J. A. (1980) "A Circumplex Model of Affect." *Journal of Personality and Social Psychology*, 39: pp. 1169-1178.
- Scarantino, Andrea (2016) "The Philosophy of Emotions and Its Impact on Affective Science." In Lisa Feldman Barrett, Michael Lewis, and Jeannette M. Haviland-Jones (eds.) *Handbook of Emotions* Fourth Edition, The Guilford Press, pp. 3-48.
- Scarantino, Andrea (2009) "Core Affect and Natural Affective Kinds." *Philosophy of Science* 76 (5): pp. 940-957.
- Schwarz, Nobert, and Gerald L. Clore. (1983) "Mood, Misattribution, and Judgments of Well-Being: Informative and Directive Functions of Affective States." *Journal of Personality and Social Psychology* 45(3): pp. 513-523.
- Sumner, L. W. (1996) *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press.
- Suzuki, Makoto. (2011) "Well-being, Desire, and the Problem of 'Miswanting': Rethinking

the Philosophical Theories of Well-being and the Practice of Informed Consent in View of Psychological Studies.” In *Applied Ethics: Old Wine in New Bottles?* Center for Applied Ethics and Philosophy, Hokkaido University, Japan, pp. 13-22.

Tappolet, Christine. (2015) “Value and Emotions.” In *The Oxford Handbook of Value Theory*, ed. by Iwao Hirose and Jonas Olson, Oxford University Press. pp. 85-95.

Zimmerman, Michael J. (2011) “Partiality and Intrinsic Value.” *Mind* 120: pp. 447-483.

戸田山和久 (2016) 『恐怖の哲学—ホラーで人間を読む』 NTT 出版.

信原幸弘 (2017) 『情動の哲学入門—価値・道徳・生きる意味』 勁草書房.

廣中直行 (2003) 『快樂の脳科学—「いい気持ち」はどこから生まれるか』 日本放送出版協会.