

第十七回ワークショップ（論文草稿検討会）

2012年7月18日開催

■参考資料

井頭 昌彦 2012. 「「心を持ったロボットをつくる」というプロジェクトはどのようなものでありうるか？」 『思索』 45(1): 389-418.

■報告者による補足説明（※当日ハンドアウトは本先端研HPトップページから入手可）

I 「心を持ったロボットをつくる」というプロジェクトの背景

i. プロジェクトの目的：心を持ったロボットを作る

＝ロボット工学における抜本的革新の探求

ii. 問題：機械であるロボットが心を持つことが可能なのか？

どのような機能を搭載すればロボットが心を持つと言えるのか？

→ 心というものを我々が明確な仕方で理解できていないことがネック

↓

iii. 着目点／アプローチ：「心とは何か」という哲学的問題に焦点化

→ ロボット工学と哲学の連携というプロジェクトの構想へ

II 原稿内で展開したメタ哲学的議論について

i. 方法論的指針：心の哲学から理論評価規準として「直観と理論の合致」を抽出

ii. 「心を持つロボット」を作るために提示したルート：

「フルスペックの心」から「痛みを感じられるロボット」へ課題をダウングレードし、「痛みというのは生き物が感じるものだ」という直観を覆すためのルートを提示

①解釈主義的理解をベースに痛みを理解する。

②そのうえで、刺激やふるまいのパターンだけでなく、道徳的配慮の対象であるということも「痛み」の概念の理解の中に組み込み、解釈主義的理解をブラッシュアップする。

③実装すべき機能を絞り込み、チートの可能性を縮減する二つのタスクの達成を目指す。

④タスク達成により作られたロボットの性能が良ければ痛み＝生き物直観を覆すかもしれない。反対に、痛み＝生き物直観が強くと、ロボットの性能が良くても直観が覆らない場合には、次世代教育という奥の手があるだろう。

1. 基本事項の確認

1-1. 「痛み」の言語実践の階層的発展モデルについて

第一段階の間主観的状況における「痛み」概念の使用のレベルでは、「痛みを感じられるロボット」を作ることが可能かもしれないですが、その認定が、第二、第三段階まで敷衍されるだろうという主張を正当化する根拠はどこにあるのでしょうか？（大杉）

▶ 今回のプロジェクトの目的は、第一段階まで達成でき、認められればOKです。（井頭）

では、痛みが主観的であるという理解や神経的なものだという理解（第2、第3段階の直観）は消去されるということですか？ でも、第二、第三段階と関連する直観を私たちは実際に持っているのではないですか？（大杉）

▶ 各々の段階は別々の判断規準を持っているので、第一段階の結果が別の段階の規準を覆すことにはならない。各段階を区別するポイントは、痛みの有無を判断するうえで“決定的な”条件・規準が違うということです。痛みを主観的定義に定位して理解しても、そこで平行して神経科学的状態が起っていることを否定することにはなりません。（井頭）

それは要するに科学实在論の立場に立たないという前提のうえでの議論だと思うのですが、でも、第二、第三段階では——实在論が科学の世界だけにあるものではないので——实在論的な直観も含まれるわけですよ。するとこのプロジェクトの成功は、最終的にそれらの实在論的直観を否定することになりませんか？（大杉）

▶ そうですね。このプロジェクトがもし成功すれば、ロボットが主観的状態や神経系を持たないにもかかわらず痛みを感じていることが認められることになるため、公共的に観察可能なものの背後や向こう側に感じられる「痛み」があるという直観は、弱められるか、根拠がないということになるでしょう。（井頭）

1-2. 志向的スタンスについて

注21の記述とも関連しますが、デネットの志向的スタンスというのはどのようにして客観性を認められるんですか？（大杉）

▶ ここで言う客観性というのは、われわれの理解の“向こう側”の話ではなくて、わりと弱めの客観性なんです。たとえば志向的スタンスというのは、お父さんが冷蔵庫を開けてビールを取り出すという場合、お父さんがビールを飲みたかったんだという欲求を帰属させ、かつそこに冷蔵庫にはビールがあるはずだという信念を帰属させることによってうまくその行動を説明できるのであれば、お父さんは信念のような志向的状态を持ちうる存在者として認められるし、その説明が成立する段階で、そういった信念を持っているということになる。デネットがそこで客観性の話をどう使っているかという、実際にそのような解釈をする人が現実には

いなかったとしても原理的にその解釈をうまく行うためのリソースがあるなら、そのシステムは志向的システムであり、その存在者は信念、欲求システムを持っていると言ってよいという仕方で用いている。つまり、実際に解釈されるかどうかではなく、解釈可能であるかどうかが大前提という考え方なんです。ここで客観性を確保している。われわれが解釈するか否かに従い、信念や欲求というものが生じてくるようなものではなくて、われわれが実際に解釈するか否かに関わらずそこにあり、かつ、われわれが解釈することによってそれが明らかになるという、そういう微妙な考えをデネットはとっているんですね。

で、現象的スタンスは基本的にはこのような解釈ができるものなんですけれども、僕の今回の原稿では、実際に、人間の側に痛みを配慮する実践が生まれたことをもって、そこに痛みを感じられる機能や能力があることを帰属させるという使い方をしてるので、そういった解釈がとれないような使い方をしてる。そうすると、アクチュアルな解釈実践がなされたことをもって痛みの生起と見なすということになりかねない。ここを上手く整理しなくてはならないというのが、注21のポイントです。（井頭）

どうして志向的スタンスのようにできないんですか。アクチュアルな解釈実践が行われているかどうかを、実験では証明できないということですか。（大杉）

▶このプロジェクトを進める都合上、そういう制約を設けざるを得ないということです。（井頭）

1-3. 直観的理解を変えるという戦略について

一番初めの問題（1-1）とも関係すると思うのですが、どうしても直観的な理解と食い違う場合は直観の方を変えて行けばよいということでしたが、理論と直観の双方が可変的であるとしたら、どのような理論も認められるということになりませんか？（大杉）

▶これは非常に重要な問題です。基本的には、直観を変えることはすごく難しいということと、ある程度方向性を絞ってしまうことができると思います。ただ、直観がどれほど変えられるのかということについては、意見が別れるところです。この直観と理論の整合性モデルに関して、直観が絶対変えられない・変えられるべきではないというスタンスをとっている現代の哲学者も何人かいますが、その一方で「がんがん変えていけよ」と思っているローティのような人もいます。僕はちょっとローティ寄りで、そうすることでメリットがある場合には能動的に変えることもいいんじゃないかという立場です。メリットをちゃんと提示して、直観の部分を変えることでかかるコストを相殺するだけのメリットを提示することによって初めて、直観を変えることが認められる場合もあると僕は考えます。（井頭）

メリットという違う評価基準が出てくるんですね。ローティが入ってくるんですね、そこで。プラグマティックな、改良主義みたいな。（大杉）

▶ローティは役に立つという評価基準を採っているというよりも、役に立つかどうかということにそんなに重点を置かずに、とにかく増やせ、というファイヤーベント的なスタンスをとっているところがある。そういう時のローティは、おそらくもっと過激な増殖主義者でしょう。（井頭）

2 道徳的な配慮実践とはどのようなものか？

2-1. 擬人、擬人化とは何を意味するのか？

「擬人化ではなく、真正の道徳的配慮の対象として」という言い方がなされていたりするわけですが、擬人、あるいは擬人化とは、どういうことを意味するんですか？（深澤）

▶たとえば、自動的に動くぬいぐるみがあったとします。それに対して子どもがボールをぶつけて遊んでいたとする。それを見た母親が「かわいそうでしょ、やめなさい」と指導し、子どもがぬいぐるみを「よしよし」と撫でるとする。この時、表面上、ぬいぐるみにたいして道徳的配慮がなされているように見えると思うのですが、その後子どもがそのぬいぐるみをおもちゃのように放り投げるということを多分やると思うんですね。その時、道徳的配慮をしていたように見えたそのふるまいは整合性という観点からおそらく擬人化して道徳的配慮の対象として扱うかのようなふるまいをしていたということになると思う。そうしたものを排除したいということです。（井頭）

本当に持続的に、配慮し続けると。（深澤）

▶そうですね、ここは結構致命的なところだと思っていて、明確な基準を設定することは多分難しいと思うのですが。なるべくそういう形で理解して頂ければ。（井頭）

2-2. チートの可能性の排除について

定義上、配慮実践がある場合、チートの可能性が排除されているという記述がある一方で、配慮実践を作り込むことでチートの可能性を縮減するという記述がある、その整合性はどうか？また、人工物であることが分かってしまった段階で、チートの可能性は排除できないのではないのでしょうか？（深澤）

▶後者の方から言えば、痛みは生き物しか持つことができないのだという直観と似たような位置づけにあるわけですね。これを人工物であっても痛みを持つものとして扱わざるを得ないだろうという状況を作り出せば、直観の方を変えることも出来ると思うので、そうした仕方に対処することを目論んでいるという回答になると思います。前者の方に関しては、完璧に道徳的配慮実践を惹起するものができたら、その段階で完璧に道徳的配慮の対象にしているんで、チートではないと。しかし実際に作るようになった場合には、道徳的配慮実践をどれくらい惹起できるかということは多分程度の問題になってくると思います。プロジェクトのプロセスとゴールの違いがご指摘された記述の違いに重なっているということですね。（井頭）

実験状況をコントロールするという前提があるのでそれはすごくよく分かりますが、現実実践を行ってみた場合には、複雑で両義的な反応があり得ますよね。たとえば配慮実践を含むより広い共感実践のようなものは、身体とかそういうものにたいする態度と同じで、信じているかどうかという問題を含めて、両義的な反応があり得ると思うんです。先ほどのぬいぐるみ

の例では擬人化と持続的な配慮という区別が提示されてましたけれど、擬人化もその時々にはすごく配慮しているというか。（深澤）

▶そこももう一つ致命的なところなんです、それには配慮の種類を分類するということで応えられる。ただしふるまいのレベルだけで識別することには限界があつて、最終的には「壊れちゃうでしょ」と怒るか、「可哀想でしょ」と怒るかという、そういう言語化された仕方
で配慮実践を分類するやり方しかないだろうと僕自身は思います。擬人化のケースと道徳的配慮のケースの区別が曖昧だという問題については、現在僕の手元にはそれを識別するリソースがないので、実験状況におかれた人のふるまいや発話のパターンを分析してもらう人を用意するという対応をとることになると思います（社会心理学者、エスノメソドロジスト、神経科学・認知科学・脳科学的者など）。（井頭）

2-3. 道徳的配慮を行うこと＝心を持つこと？

道徳的配慮とは何なのかということと関わる質問なんです。たとえば、虐げられる対象が心を持っている現象的システムであるという認定のもとに殴るサディストを例に考えてみると分かりやすいのですが、道徳的配慮を行うということが配慮対象が心を持つということすべて覆っているということにはならないですね。その場合、なぜ道徳的配慮が特権的に扱われるのでしょうか。要するに、私たちのネガティブな反応やポジティブな反応に関係なく、もっと中立的な配慮のあり方があるような気がするのですが。（大河内）

▶もしそのような中立的な配慮のようなものがあれば、非常にいいなと思います。まず、道徳的配慮を特権視しているわけではなくて、とりあえずそれしかないからこれを扱っているというのが一つ。もう一つは道徳的配慮の実践があるかないかという基準で心を持っているかを識別すると、削ぎ落とされてしまう人たちが出てくるとは思うけれど、今それは問題ではない。道徳的配慮を持たない人たちが正当に扱うということを目的としているわけではないので、とにかくロボットであるにもかかわらず痛みを感じられるという認定を引き出すという目的に関しては、そういうことは問題とはならないと思います。（井頭）

3. 「痛み」に注目することの正当性について

とりあえず戦略的に扱いやすいから痛みというものを基準としたのだと思うんですけど、痛みを持つということと心を持つということの関係をどう考えるのでしょうか。少なくとも、痛みを持つ存在（e.g. 動物）であることと心を持つ存在であるということは一致しないと思いますが、その時に痛みを持ち出して、心を持ったロボットを作ることについて論じることが適切なのでしょうか？（大河内）

▶すごくきつい質問だと思います。とりあえず原稿では、痛みを感じられるロボットを作ってみてある程度満足のゆくものが完成した暁にそれが心を持っているものとして認められるようだったら、別の側面として思考や意識を作り込むことについても考えてみようという話をしたと思うんですけど、この話にはどこか変なところがあるのは確かです。痛みを持つロボットが心を持っていると言えるという条件が特定されていないので、だから変なんだと思うんですけど。どうすればいいか分からないんですね。多分これは、心の哲学だけではなくて心とい

うものに関する幅広い議論を参照しながら、心に関する概念連関を持つ何らかの公共的な事柄をひっばってくる必要があると思いますが、現時点では何をひっばってくればいいのか全然想像がつかないので、一応そこはオープンにしています。（井頭）

4. 研究の意義づけ

4-1. 解釈主義の適用について——解釈主義の難点は回避できるのか？

原稿の初めに、心身因果を説明できないという解釈主義の抱える難点について指摘されていますよね。ロボットの構想のところで解釈主義を使用しているのはそれが正しいからというよりもアプローチとして使いやすいからだと思うんですけど、その場合、最初に指摘されている難点は残っていますよね。それもオープンにしておく？（大河内）

▶僕は、解釈主義について現状で指摘されている難点を踏まえたうえで、その難点を回避するやり方があるだろうと実は思っている。解釈主義が心身因果を説明できない一つの理由というのは、物理主義を前提とするからなんですね。物理主義を前提としてしまうと物理的な身体運動をひき起したのは別の物理的な身体運動であるということになるので、心が何かをひき起したことはない。そのため解釈主義のアプローチをとったとしてもそれは物理的因果関係に、因果的効力を持たないものを重ね合わせるだけにしかならないのだという反論があるんですけども、それは物理主義を前提とするからダメなのであって、物理主義を前提としない仕方でも因果関係を捉えることができれば全然問題がないと僕自身は考えています。（井頭）

そうなるかと全く別の議論がまた必要になるということですよ。解釈主義って要するに、基本的に外から心のようなものを帰属させるという議論ですよ。その存在ないシステムが全く物理的に動いているとしても帰属させられればよいという議論ですよ。（大河内）

▶はい。そうですね。解釈主義と物理主義の相性の悪さというのは機能主義を例にすると分かりやすいんですが、機能主義の場合には特定の欲求を持っているということが特定の神経科学的状态と同一視されるんですよ。そうすると心的状態と物理的状态が同一視される。で、解釈主義の場合にはそういう風にできない。なぜかという解釈の不確定性があるから。これにより内容に関する同一性が物理主義の枠内ではキープできなくなることも、解釈主義の難点の一つなんです。これは物理主義的一元論をとらないで多元論で行きましょうという形で、回避できる。（井頭）

4-2. 物理主義は覆すことができるのか？

それでさっきの話とも関連するのですが、物理主義的な考え方は理論だけでなく極めて持続性を持った形で直観のなかにも刷り込まれていると思うんですよ。そこでまず、直観の変化するスパンや時間のリズムと工学・技術者たちの時間のリズムというのは明らかに違うだろうと思うわけです。そのうえで、物理主義がまた直観のなかに深く根付いているとすると、やはりロボットを作ってみても、そんなに容易に物理主義が覆されるようには思えない。つまり、物理主義の問題って直観にも浸透しているような非常に大きな問題で、心を作ることがで

きるかという問題に付随するものとして扱えるような問題ではないんじゃないかと思うのですが。（大杉）

▶まず、直観の変化のスパンと工学的・技術的な取り組みのスパンが全然違うという話についてなんですが、今までの歴史を見ると基本的にそうだと思うんですね。しかし、それは、こういう構造があるということに気づいていなかったが故に能動的にそういった変化を起こせなかったから、そうだったに過ぎなかったのかもしれない。つまり容易ではないんだけど、変えられないと思われていることも仕組みを変えることで直観は変えられるということを示そうという、すごくチャレンジングな課題ではあります。二つめの指摘については、最終的に僕が物理主義者を説得しきれるというところまで目指すというのであれば確かにそうです。でもそのためのアイデアを示しておく、その道筋をつけることについてはこの研究にはおそらく十分な意味があると思います。（井頭）

4.3. 直観を変えるというアプローチについて——これは哲学の論文なのか？

僕の印象としては、心を持ったロボットを作るという話が、いつの間にかどういふ風に私たちの信念を変えることができるかという話になってしまっているような気がして。それは先にそういうロボットを作ってしまうと、ロボットが心を持っていると言ってもよいだろうという論理ですね。それはもう哲学というより科学の論文なのかな、と思うのですが（別に、井頭さんの論文を攻撃しているわけではないのですが）。（大河内）

▶直観を変えるというやり方で心に関する理論の正当性を高めようというアプローチにたいする警戒心というのは、すごく良く分かる。ただ、そういう警戒心を示す人たちがそれで何をしていることになるのかというと、彼らはわれわれの概念理解の範囲内でロボットが心を持てるのかどうかという問いに答えろという要求をしているんだと思うんですね。でも、われわれのこの概念理解自体、歴史的に構築されてきたものでしかないわけですし、何らかの根拠があってそういう風に形成されてきたものでもないわけなので、この概念理解に拘泥しなくちゃいけないという枠組は、そんなに強くキープしなくても良いのではないかと僕自身は思うんです。

もう一つは、哲学の議論かという疑問についてですが。概念やそのメカニズムについての問題を扱っているのが哲学の議論であるといっても良いとか、哲学であるための条件を明記して下さいというように、この批判に対してはいくつか対応の仕方があると思う。今一番気に入っているのがp. 26注32に書いてある通りで、「この報告を哲学系の学会で聞きたいですか？」と訊いて「Yes」の返事をもらえるなら、これは十分哲学的であると言ってもいいだろうというものです。（井頭）

ある問題が哲学的であるかどうかについて研究者同士が話しているとして——「メタ哲学」ですね——、でもそこで「今僕たちが話しているんだから、それは哲学だよ」というのは根拠にはならないですね。もう少し違う根拠を出してもらえないと。（大河内）

▶なるほど。その場所が成立している段階で、メタ哲学的な新しい議論をしていることは認定してはもらえているわけですね。であれば、そういった意味で哲学的……（井頭）

一点目について補足すると、僕は直観が絶対守られなければならないとは考えていなくて、むしろ、心の哲学の議論で「これは直観に反するから反証された」という論法は、すごく疑わしいと思っています。わたしたちが持っている直観のどれかを否定するという時に、一番ストレートなやり方というのは、その直観が間違っていることを証明することだと思うんです。でも井頭さんが今やっていることはそうではなく、こっちの方がこの議論にとっては得だからそっちで行きましょうということをやっていると思う。（大河内）

➡はい。まず、二段階で考えていて。第一段階としてはふるまいや刺激のパターンを上手く組み込み、かつ道徳的配慮の対象として配慮実践を惹起するようなロボットを作る。その上で、ロボットが痛みを感じているという認定がなされたら、そこにこのロボットは痛みを感じられるという直観が一つ付け加わるんだと思うんですね。すると生き物しか心を持たないという直観と対立が生じる。そこで「生き物＝心」直観が弱ければ、この直観が間違っていることになり、直観の配置が変わるだろうと。この場合には、大河内さんがおっしゃったようなストレートなストーリーに乗ると思う。この場合には問題がない。で、もし「生き物＝心」直観がすごく強ければ、それを変えるために次世代を教育しようという話になると思うんです。僕がやりたいことの半分は、まずわれわれの理論空間の中にあるこの直観の身分がどのようなものかについて問い直したいというのが一つ。その点に関しては僕はメタ哲学的な主張をしているつもりなので全く問題はないと思うのですが、問題は実際に教育によって直観を変えようという方だということですよ。（井頭）

5. 直観的理解について

5-1. 本論文における直観の論理的身分について

途中までは直観が検証基盤であったのが、あるところから操作対象になりますよね。つまり、極端なことをいえば、直観の検証基盤とするという意味そのものがなくなってしまうように思うんです。この論文のなかでの直観の身分というか、それをどういう風に理解したらいいのでしょうか？（深澤）

➡この論文のなかで役割が変わっているように見えるというのは、その通りです。評価規準としての直観の性格と関わってくるんですが、正当化基準として用いるというのはもちろんそうなのですが、痛みに関する議論の継続性、意味内容の同一性をキープするためにも直観の一定程度の固定性は必要なんです。たとえば、痛みに関する7個の直観があってそのうち6個を捨ててしまったら、これはもはや痛みに関する事柄ではなくなってしまいます。同じ理由で、全部が変えられないということではなくて部分的に変えられるということはもちろんあり得る。だから、なぜこれを変えなければいけないのかという理由を、意味の同一性、われわれの概念理解の接続可能性のようなものに置くのであれば、直観を変えられる操作の対象とすることについては、そんなに問題があるということにはならないという答えになると思います。（井頭）

直観を操作対象にすること自体には問題ないと思うのですが…そうですね、あとそのどれを操作対象とするとか、そういうことはどのように選択されるんでしょう？最初から選択されたもので議論が進んでいる？（深澤）

▶今日配布した資料では特定していますが、原稿の方ではオープンにしている。実は、どういう仕方でそうした選択がなされるかについてのイメージが全くないんです。ただ製作・実験の過程で、どういう仕組みで何が選択されるのかについて調べるための手掛かりを得られると思う。そうすることで概念ネットワークの改訂メカニズムのようなものを探る手掛かりになる。だから、どの直観を改訂対象として選択するのかを決めて行くというのは、研究の課題でもあります。（井頭）

概念ネットワークの改訂、概念エンジニアリングというのは操作対象としてということですよ。今の話で言うと概念ディスクリプションというのかな、そのディスクリプションからエンジニアリングへの移行が、どういう動機というか正当化というか、そういうプロセスを経ているのかというのを、聞いたかったのかもしれない。（深澤）

▶ここに心に関する一つの直観があるとして、ロボットにたいする配慮実践が上手く惹起された段階で、そこにもう一つの直観が付け加わります。この別の直観を割り込ませるという作業がエンジニアリングの一部になっていると思うんですよ。そうした時に、何が付け加わって何が脱落するのかを見て行くというのが概念ディスクリプションだと思うんですけど。で、何がどうなるかが分からなくてもこうした介入実践を行って、似たようなアウトプットを引き出すことができると思うんです。だから、起こった現象を見ながらまた記述の内容を微細にして行くという。（井頭）

じゃあ、ディスクリプションとエンジニアリングというのは移行というより、同時的で入り組んでいると、そういう風に考えればよい？（深澤）

▶そうです。（井頭）

5-2. 直観をどのようなものとしてイメージすれば良いか？

直観という言葉は、この種の議論では、一般的に使われる言葉なんですか？（大杉）

▶もともとかどうかは分かりませんが、J. ロールズの『正義論』で使われている言葉なんです。道徳的な理論と道徳的な直観を反照的に使いながら整合性を取り続けて行くというようなやり方が彼の倫理学の方法論として提示されている。そこで既に直観という言葉は使われている。1970年代くらいから標準化されているようです。かつ最近では、哲学理論の正当化の基盤となりうる一般の人びとの直観というものがどういうものかということが哲学的方法論などでしばしば議論される。メタ哲学の議論のなかでは、直観と理論という対比は標準化された構図になっています。（井頭）

それほど標準化されているのだったら、それなりに直観という言葉を使う意義もあるのかもしれないんですが、脱文脈化や文脈化ということとあえて近付けて言うと、直観を文脈に含

みうるかという風に言えることだと思うんだけど。直観というのは、その言葉が喚起するような、どこからともなく到来する根拠のないものというものではおそらくなくて、直観自体がまさしく解釈主義的な取り組みを必要とするような心をめぐる様々な取り組みの運用によって支えられているわけですね。だから（プロジェクター資料で）図式化されたようなドット状の、何か一つの単体として存在しているようなものではなくて、直観そのものが網の目のようになっていくというイメージがあって、しかもそれが科学的な取り組みを含めた様々な人間の取り組みによって支えられていると考えると、この文脈のなかのこの直観を操作したり変更を加えたりということは、なかなか言えない気がするんです。（大杉）

➡そのこと自体には全く同意できると思っています。痛みは生き物が持つものだという直観を単品で括り出すことは多分できない。これを変えた段階で他のものが一つと変わってくるはずですし。ですので、これはすごく模式的な図だと理解して頂きたい。ただその一方で、この直観が否定されたという言い方自体はキープできると思う。痛みは生き物が持つものであるという直観自体が変わったということは、それだけが変わったわけではないけれども、言うことができる。なぜなら、ロボットは生き物ではないということとロボットは痛みを持つということが組み込まれていたの、そのことは論理的に両立しないからですね。だから、これが変わったということは言えるわけだけれども、これが変わったことと共に他のものも変わったということについては合意できるわけです。（井頭）

5-3. 哲学理論と日常生活における「実在論」と「反実在論」

実験の場面では、被験者が配慮実践を行ったという行為自体が、心の実在論的な主観的感覚論を再強化するように作用するといった可能性は考慮していないんですか？僕の想像では、実験に参加している被験者たちは、色々なものに支えられている機序のなかで「ロボットが可哀想だと私が思ったので配慮した」と自分の行為を理解するんだと思うんですね。そのことはむしろ、井頭さんが否定しようとする当の心の実在論的直観というのを再強化するようにも作用しますよね。（大杉）

➡もしかしたら僕は、大杉先生のおっしゃる実在論的な直観というものを上手く理解できていないのかもしれない、と今思ったんですが。（井頭）

ああ、つまりこの実験は、単にロボットに心が認められたという実験ではなくて、反実在論を支持する主張に繋がっている実験だと思うのですが、ただ、そのロボットに反応する人間がいるわけですね。つまり第三者はそれこそ行動主義とか解釈主義の枠組を当てはめればいいけれども、そこに参加する被験者は実在論的な前提に基づいて、行為するんじゃないですか。つまり参加している人も様々な実在論的直観に支えられてその実験に参加していて、しかもこの実験でなされていることは、実在論を再強化してしまうような実験でもあるんじゃないかと思うのですが。（大杉）

➡実験が実在論を再強化する可能性については全然考えていなかったです。もしかしたら再強化するかもしれないですけど、とりあえず分けて考えるべきことがあると思っています。痛み感覚に対して配慮実践をするという形でのふるまいのレベルで着陸させて理解することと、行為を何らかの配慮実践だとか公共的な実践に着陸させるものとして理解することは、

心的なものを理解するという意味では共通しているわけなんですけれども別の事柄なので、一方の話がもう一方の話に影響を与えとは考えていない。僕自身は行為に関しては、行為をどのように記述するかということが行為の識別に関して決定的に重要な役割を果たすと考えているので、もしそれが実在論的な想定だということであれば、それを引き受けます。

実在論とか反実在論をどう定義づけるかというのは、僕が原稿のなかで明確にはしていないのが良くないんですけれども……とりあえず、ここでやっている話は、われわれがそのように理解・記述するというその仕方が全部間違っているという可能性が認められるかどうかで実在論か反実在論かを分けた方がいいように思う。つまりわれわれがみな痛みを認定するからといって本当に痛みがあるかどうかは分かりませんよね、という想定を持ち込むのが実在論者で、われわれが痛みを認定することそれこそが痛みがあるということなのだ、とその“向こう側”のようなものを考えないのが反実在論者というように。このように理解して頂けると、被験者のレベルでリアルなものとして痛みを認めるとしても反実在論の内部で成立する話だと思うんです。（井頭）

僕が聞きたいのは、理論とその正当化の話ではなくて、直観の話とその変更に関する話なんです。（大杉）

学習段階モデルの第二段階、第三段階にあたることを被験者がやっているということですか？つまり、被験者がこのモデルとは逆に間主観的实践に先立って、初めに自分の主観的感觉に基づいた仕方で配慮実践を行っているのではないかということですか？（大河内）

そう、実験そのものがその直観をさらに深めるじゃないですか？（大杉）

今の大河内さんの説明が正しいのだとすると大杉先生の疑問点は、解釈主義を前提としたうえで組まれた実験であるはずなのに、そこで行われている人びとの実践自体が解釈主義から離れた内観主義的な想定の方を強化してしまうのではないかということにあるんですかね？（井頭）

想定、いや直観を、ですね。理論家ではないですから、被験者は。だから、その実験は直観を変えらるということには肯定的に作用しないのではないかと思います。この実験もまた長らく維持されてきたような実在論と反実在論を、うまく両方とも生み出し続ける。いや、科学的理論を直観とってしまふと分かりづらけれど、要するにその直観を成り立たせている制度的・社会的なネクサスと、科学的実践のそれも等価なわけですよ。だから、さっき脱文脈化と文脈化という言葉で言ったことと関連して(5.2)、理論も直観も両方とも互いに連なりながら成り立っている織物だから、実験もそのカレイドスコープ的な織物の一部になってしまうという、そういうイメージですけれども。（大杉）

なるほど。いや、まだ十分に消化できていないですけども、すごく痛いところをつかれてるんじゃないかという気がしてきました。要するに僕が、解釈主義的な痛み理解というのをベースにして色々な実験を組んできているわけですけども、それを完遂できていないんだという話になるんだと思うんですよ。解釈主義的な痛み理解、この直観の定義というのがどこかで崩れてしまうような話が入ってくると。僕は初め、この話を通して、解釈主義的な心の定義というものを間接的に正当化しようとしていたんですよ。でも、そのなかで、そうじゃ

ない考え方が介在しないと、この正当化実践が成立しないということになれば、困っちゃいますね。（井頭）

6 痛みと心を結び付けることについて

痛みと心の結び付けの問題について伺いたいんですが。配慮実践を引き出すという意味では、痛みと心をつなぐというのは、すごく有効だと思うんですね。ただ一方で、原稿にも書いてありましたけれど、普通は、認知的キットとかね、痛くって話しかけると答えるとか、そっちの方が「心的なもの」という気もする。痛みというのは、心の現象のなかでもちょっと特殊な気がするんですね。配慮実践を引き出すという意味では、痛みに注目するのは巧妙だなと思ったんです。でも実機を作るという場合に、心に関する類似の反応を引き出せるような、痛みとかいう前にもっと簡単な、仕掛けができるんじゃないかと思ったんですよね。（深澤）

▶心を最終目的にする場合に、思考とか想念とかではなくて、痛みを扱うのはなぜなのかということですか？（井頭）

思考とかコミュニケーション、ですね。（深澤）

▶はい。すごく重要な問題だと思います。一つは、思考に焦点をあてるとチートの可能性が増大するから、という理由があります。もう一つの理由は、思考を行っているかどうかというのは、ロボットが心が持つかどうかを判断する時に「持つ」と判断する人がまず理由にあげるようなポジティブな認定の条件になっているものなんですね。反対に、心を持つかどうかという問いに対して「持たない」と答える人たちがその判断の理由として最も多くあげるのが、感情なんです。ポジティブな理由とネガティブな理由が違うというのが、面白いところでもあるのですが。心を持ったロボットを作るためのミニマムな条件を充たすという意味では、ネガティブな要因から排除する方が戦略的にはいいと思っているということなんです。（井頭）

（完）