This chapter on the web
informationphilosopher.com/freedom

Free Will

In our 2011 book *Free Will: The Scandal in Philosophy*, our mind model was a combination of a rudimentary *experience recorder and reproducer (ERR)*[1] and our *two-stage model of free will*. Recent information analysis of the mind and the mind-body problem has greatly strengthened our mind model. We now see the mind as *immaterial* information, the "software in the hardware" of the material brain, which we view as a *biological information processor*.

Five years ago, we saw the quantum randomness in the first stage as adding "uncaused" events to fit a picture of "*event causality*" and to attack the "causal closure" of the eliminative materialists.

Now that our mind model is unapologetically *immaterial*, it is in fact an example of the kind of *metaphysical* entities that the famous philosopher P. F. Strawson rejected as "panicky metaphysics - uncaused causes, immaterial minds, non-empirical noumenal selves, non-event agent causes, and prime movers unmoved." We now endorse the idea of *agent causality*,[2] in which the mind has causal powers over the material world.

We argue that freedom of the will begins in the pre-deliberative *thoughts* of the agent. Although Albert Einstein was a strong believer in determinism, he saw our thoughts and theories as "free creations of the human mind." These creative thoughts bring *new information* into the universe. New information *emerges*[3] from the material and biological worlds to become part of the mental or ideal world, even as it is embodied in the material world.

Without alternative possibilities for an open future, there can be no new information in the universe, in biology, or in human minds. But there continues to be new information, in stars still forming, in the evolution of new species, and in creative minds.

1 See Appendix E.
2 See informationphilosopher.com/freedom/agent-causality.html
3 See chapter 27 on emergence and appendix F on cosmic creation..

The Two-Stage Model of Free Will

Our two-stage model is now the most plausible explanation, not only for human free will, but also for *creativity*, cited in the American Psychological Association's *Review of General Psychology* as supporting the Campbell-Simonton BVSR model of creative thought.[4]

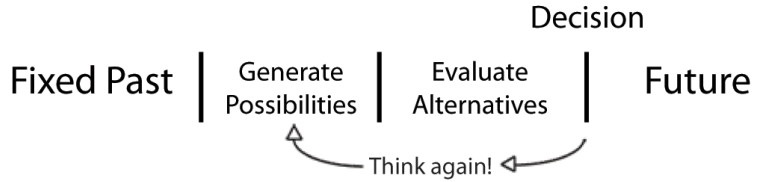
Given the "laws of nature" and the "fixed past" just before a decision, many philosophers wonder how a free agent can have any possible alternatives. This is partly because they imagine a timeline for the decision that shrinks the decision process to a single moment.

Collapsing the decision to a single moment between the closed fixed past and the open ambiguous future makes it difficult to see the role of free thoughts of the mind - which bring new information into the universe - followed by the willed and adequately determined action in a temporal sequence, as shown here.

But the two-stage model is not limited to a single step of generating alternative possibilities followed by a single step of determination by the will. It is better understood as a continuous process of possibilities generation by what we call the micro mind (parts of the brain that leave themselves open to noise) and adequately determined choices made from time to time by the macro mind (the same brain parts, perhaps, but now averaging over and filtering out the noisiness that might otherwise make the determination random).

4 Review of General Psychology, APA, 2013, Vol 17, No 4, 374

In particular, note that a special kind of decision might occur when the macro mind finds that none of the current options are good enough for the agent's character and values to approve. The macro mind then might figuratively say to the micro mind, "Think again!" Thus we can say that the agent has control over the generation of alternative possibilities, without controlling the specific new idea that may come to mind



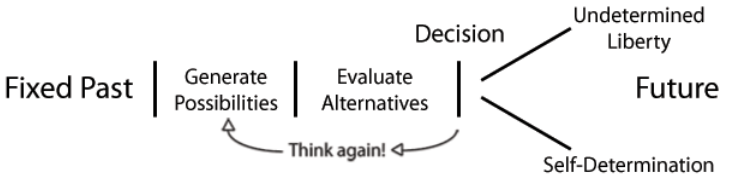
Many philosophers have puzzled how an agent could do otherwise in exactly the same prior circumstances. Since humans are intelligent organisms, and given the myriad of possible circumstances, it is impossible that an agent is ever in exactly the same circumstances. The agent's memory (stored in the *experience recorder and reproducer*) of earlier similar circumstances guarantees that.

The two-stage model may make an artificial temporal separation between micro-mind creative randomness and macro-mind deliberative evaluation. These two capabilities of the mind can clearly be going on at the same time. That can be visualized by the occasional decision to go back and think again, when the available alternatives are not good enough to satisfy the demands of the agent's character and values, or by noticing that the subconscious micro mind might be still generating possibilities while the macro mind is in the middle of evaluations.

Finally, not all decisions in the two-stage model end with an adequately determined "de-liberation" or perhaps better we can call it simply self-determination. Many times the evaluation of the possibilities produces two or more alternatives that seem more or less of equal value.

In this case, the agent may choose randomly among those alternatives, yet have very good reasons to take responsibility for whichever one is chosen. This is related to the ancient liberty of indifference.

I like to call such a decision an "undetermined liberty," because it remains undetermined at the moment of the decision. Though not determined by the deliberations, we can say that the agent "deliberately" chooses at random between equal options.



Undetermined liberties include Robert Kane's Self-Forming Actions, although Kane limits his SFAs to "torn" decisions between moral and self-interested alternatives.

Neuroscientific Evidence for the Two-Stage Model

Benjamin Libet's famous experiments are widely cited by compatibilists and determinists as showing that the decision has been made a long time before the conscious will can act. We shall interpret them as supporting the temporal sequence in the two-stage model of free will, creating new information in the first stage.

The original discovery that an electrical potential (of just a few microvolts - μV) is visible in the brain long before the appearance of conscious will was made by Kornhuber and Deecke (1964). They called it a "*Bereitschaftspotential*" or readiness potential.

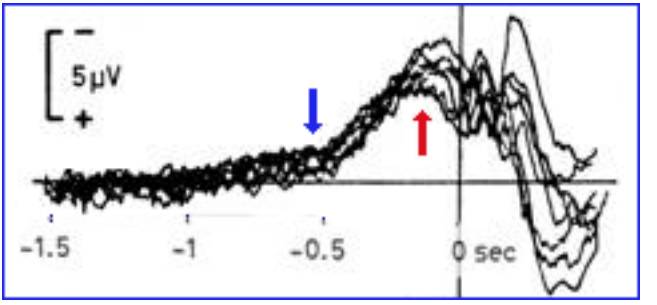


Figure 4-1. Kornhuber and Deecke "readiness potential"

The neurobiologist John Eccles had speculated that the subject must become conscious of the intention to act before the onset of this readiness potential. Benjamin Libet decided to test Eccles's idea.

Libet's 1983 experiments measured the time when the subject became consciously aware of the decision to move the finger. Libet created a dot on the screen of an oscilloscope circulating like the hand of a clock. The subject was asked to note the position of the moving dot when he/she was aware of the conscious decision to move a finger or wrist..

As shown on the RP diagram, Libet found that although conscious awareness of the decision preceded the subject's finger motion by only 200 milliseconds (the up arrow), the rise in the readiness potential was clearly visible at about 550 milliseconds before the flex of the wrist (down arrow). The subject showed unconscious activity to flex about 350 milliseconds before reporting conscious awareness of the decision to flex. Indeed an earlier very slight rise in the readiness potential can be seen as early as 1.5 seconds before the action.

Of course the kinds of deliberative and evaluative processes that are essential for free will involve much longer time periods than those studied by Libet. Nevertheless, we can correlate the beginnings of the readiness potential (350ms before Libet's "conscious will" time "W" appears) with the early stage of the two-stage model, when alternative possibilities are being generated, in part at random.

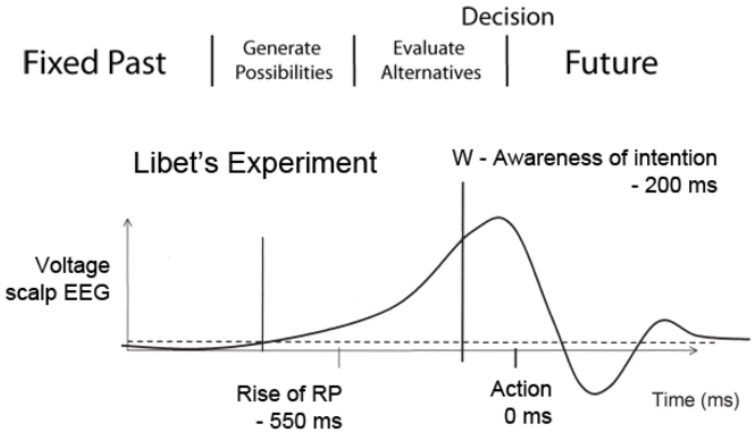


Figure 4-2. Readiness potential and the two-stage model

The early stage may be attributed to the subconscious, which is capable of considering multiple alternatives (William James'

"blooming, buzzing confusion") that would congest the low-data-rate single stream of consciousness.

Alfred Mele criticized the interpretation of the Libet results on two grounds. First, the appearance of the RP a half-second or more before the action in no way makes the RP the cause of the action. It may simply mark the beginning of forming an intention to act. In our two-stage model, it corresponds to the agent's thoughts that generate possible options, which may create new information.

Libet himself argued that even if a decison has been made, there is enough time after the W moment (a window of opportunity perhaps 50 ms) to veto the action, but Mele's second criticism points out that such examples of "free won't" would not be captured in Libet experiments, because the recording device is triggered by the action (typically flicking the wrist) itself.

Thus, although all Libet experiments ended with the wrist flicking, we are not justified in assuming that the rise of the RP (well before the moment of conscious will) is a *cause* of the wrist flicking.

Libet knew that there were very likely other times when the RP rose, but which did not lead to a flick of the wrist. All such events could create *immaterial* information about new possibilities, but might not be acted upon immediately. Libet noted that in normal decisions we might deliberate all day.

> We should also distinguish between deliberations about what choice of action to adopt (including preplanning of when to act on such a choice), and the final intention to actually "act now." One may, after all, deliberate all day about a choice but never act... However, conscious will definitely can control whether the act takes place. We may view the unconscious initiatives for voluntary actions as "burbling up" unconsciously in the brain. The conscious will then selects which of these initiatives may go forward to an action, or which ones to veto and abort so no act occurs.[5]

We conclude that Libet's neuroscientific experiments may be interpreted as supporting the two-stage model. We know little about what goes on in the early rise of the readiness potential. But only a dogmatic determinist would claim that it already contains and directly causes any later decision.

5 B. Libet, *Mind Time*, pp.148-149

History of the Free Will Problem[6]

In our research on the history of the free will problem, we have identified several thinkers who developed two-stage solutions to the classical problem of free will, first William James, then Henri Poincaré, Jacques Hadamard, Arthur Holly Compton, Karl Popper, Daniel Dennett, Henry Margenau, Robert Kane, David Sedley and Anthony Long, Roger Penrose, David Layzer, Julia Annas, Alfred Mele, John Martin Fischer, Stephen Kosslyn, Storrs McCall and E. J. Lowe, John Searle, and Martin Heisenberg.[7]

Some of course were more clear and comprehensive about the two stagesthan others, but our goal is to give them all credit.[8]

Recently we discovered a possible two-stage argument many centuries before William James.

Titus Lucretius Carus is our main source for the work of Epicurus, who provided the first argument for chance with his "swerve" of the atoms. Lucretius eloquently made Epicurus' case. Shortly after describing the swerve, he says:

> "If all motion is always one long chain, and new motion arises out of the old in order invariable, and if first-beginnings do not make by swerving a beginning of motion so as to break the decrees of fate, whence comes this free will?"[9]

But now we have found evidence that Lucretius made the case for *alternative* thoughts *coming to mind* before a willed decision, and the possible new ideas sound very much like Epicurus' random swervings.

> Now listen, and hear what things stir the mind, and learn in a few words whence these things come into the mind. In the first place I tell you that many images of things are moving about in many ways and in all directions.[10]

---

6 Doyle, 2011, chapter 7 is a 60-page history of the problem

7 See Doyle, 2011, chapter 12, for these two-stage solutions.

8 Also see informationphilosopher.com/freedom/two-stage_models.html

9 *De Rerum Natura*, Book 2, 251

10 *De Rerum Natura*, Book 4, 722

This also sounds a great deal like William James' "blooming, buzzing, confusion" of the subconscious and Libet's "burbling up."

> even in things plainly visible you can observe that, it is just as if the thing were all the while withdrawn and far removed from you. Then what wonder is it, if the mind misses everything except what it is itself intent on?[11]

Lucretius again sounds like James, who explains choice as the focusing of attention. Next comes the will (*voluntas*).

> Next I will say how it comes about that we can carry onwards our steps when we please...I say in the first place images of movement come in contact with our mind, and strike the mind, as I said before, After this comes will; for no one ever begins anything until the intelligence has first foreseen what it wills to do.[12]

So Lucretius may have long ago captured the essence of the *temporal sequence* in our two-stage model.

The classic problem of free will is to reconcile an element of freedom with the apparent determinism in a world of causes and effects, a world of events in a great causal chain.

Determinists deny any such freedom.

Compatibilists redefine freedom. Although they say that our will is determined by prior events in the causal chain (including our reasons, motives, etc.), our will is in turn causing and determining our actions. Compatibilists say that determinism of our actions by our will allows us to take moral responsibility for our actions. This is correct. The second stage of our model makes us responsible.

Libertarians think the will is free when a choice can be made that is not *pre-determined* or *necessitated* by prior *events*. The will is free when *alternative* choices could have been made with the same pre-existing conditions.

Freedom of the will allows us to say, "I could have chosen (and done) *otherwise*."

In a deterministic world, everything that happens follows ineluctably from natural or divine laws. There is but one possible future. We cannot have chosen *otherwise*,

In the more common sense view, we are free to shape our future, to be creative, to be unpredictable.

11 *De Rerum Natura*, Book 4, 815
12 *De Rerum Natura*, Book 4, 881

From the ancient Epicureans to modern quantum mechanical indeterminists, some thinkers have suggested that chance or randomness is an explanation for freedom, an explanation for the unpredictability of a free and creative act. A truly random event would break the causal chain and nullify determinism, providing room for human freedom.

Freedom of human action does require the randomness of absolute unpredictability, but if our actions are the direct consequence of a random event, we cannot feel responsible. That would be mere indeterminism, as unsatisfactory as determinism.

Moreover, indeterminism appears to threaten reason itself, which seems to require certainty and causality to establish truth, knowledge, and the laws of nature.

Most philosophers in all ages have been committed to one or more of the dogmas of determinism,[13] refusing to admit any indeterminism or chance. Aristotle said chance was "obscure to human reason." Chryssipus described the case of "indeterminism is true" as a disaster for reason. David Hume found "no medium betwixt chance and necessity." Many theologians thought chance *atheistic*, doubting God's omniscience,

Many scientists agree that science is predicated on strict causality and predictability, without which science itself, considered as the search for causal laws, would be impossible.

For those scientists, laws of nature would not be "laws" if they were only statistical and probabilistic. Sadly for them, all laws of nature turn out to be thoroughly statistical and our predictions merely probable, though with probabilities approaching certainty. Science is irreducibly statistical.

But fortunately, for large objects the departure from deterministic laws is unobservable. Probabilities become indistinguishable from certainties, and we can show there is an "adequate (or statistical) determinism"[14]

Important elements of the model have been proposed by many philosophers since Aristotle, the first indeterminist. A number of modern philosophers and scientists, have proposed models of free

13 See Doyle, 2011, chapter 9, for a review of many determinisms.

14 See informationphilosopher.com/freedom/adequate_determinism.html

will. But none of them has been able to locate the randomness so as to make free will "intelligible," as libertarian Robert Kane puts it.

The insoluble problem for previous free-will models has been to explain how a random event in the brain can be timed and located - perfectly synchronized! - so as to be relevant to a specific decision. The answer is it cannot be, for the simple reason that quantum events are totally unpredictable.

The two-stage model is not a single random event, one per decision, but many random events in the brain as a result of ever-present noise, both quantum and thermal noise, inherent in any information storage and communication system.

The mind, like all biological systems, has evolved in the presence of constant noise and is able to ignore that noise, unless the noise provides a significant competitive advantage, which it clearly does as the basis for freedom and for creativity that brings new information into the universe.

Let's see how randomness in the two-stage model is never the direct cause of our decisions. Decisions are always adequately, i.e., statistically, but near certainly, determined by reasons and motives.

We assume that there are always many contributing causes for any event, and in particular for a mental decision. In both the Newell-Simon "Blackboard" model[15] and Bernard Baars' "Theater of Consciousness" and "Global Workspace" models,[16] there are many competing possibilities for our next thought or action. Where do they come from? And, most importantly, does the agent have any control over their generation?

Each of these possibilities is the result of a sequence of events that goes back in an assumed causal chain until its beginning in an uncaused event. Aristotle called this original event an *arche* (ἀρχῆ), one whose major contributing cause (or causes) was itself uncaused.

What this means is that tracing any particular sequence of events back in time will come to one event - a "starting point" or "fresh start" - Aristotle's origin or *arche* - the dreaded "*causa sui*." Today we say it must involve quantum indeterminacy.

15 Newell and Simon, 1972
16 Baars, 1997

Whether a particular thing happens, says Aristotle, may depend on a series of causes that

> "goes back to some starting-point, which does not go back to something else. This, therefore, will be the starting-point of the fortuitous, and nothing else is the cause of its generation."[17]

We can thus in principle assign times, or ages, to the starting points of the contributing causes of a decision. Some of these may in fact go back before the birth of an agent, hereditary causes for example. To the extent that such random causes adequately determine an action, we can understand why hard determinists think that the agent has no control over such actions. Of course if we can always opt out of an action at the last moment, so we retain control, even if the origin of the option was inherited.

Other contributing causes may be traceable back to environmental and developmental events, perhaps education, perhaps simply life experiences, that were "character-forming" events. These and genetic or hereditary causes would now be present in the mind of the agent as fixed habits, with a very high probability of "adequately determining" the agent's actions in many situations.

But other contributing causes of a specific option may have been undetermined up to the very near past, even fractions of a second before an important decision. The causal chains for these contributing causes may originate in the noisy brain. They include the free generation of new *alternative possibilities* for thought or action during the agent's deliberations. They fit Aristotle's criteria for causes that "depend on us" (ἐφ' ἡμῖν) and originate "within us" (ἐν ἡμῖν).

Causes with these most recent starting points are the fundamental reason why an agent can do *otherwise* in what are essentially (up to that starting point) the same circumstances.

These alternatives are likely generated from our internal knowledge of practical possibilities based on our past experience. They are stored in our experience recorder and reproducer. Those that are handed up by the ERR for consideration to Baars' "executive function" in his "Theater of Consciousness" may be filtered to some extent by unconscious processes to be "within reason." They likely consist of random variations of similar actions willed many times in the past.

17 Metaphysics Book VI 1027b12-14

Note that the evaluation and selection of one of these possibilities by the will is as deterministic and causal a process as anything that a determinist or compatibilist could ask for, consistent with our current knowledge of the physical world.

Remember also that instead of strict causal determinism, the world offers only adequate (or statistical) determinism, and it is the random origins of possibilities that provides libertarian freedom of thought and adequately determined but not pre-determined action.

Why have philosophers been unable for millennia to accept the common sense view that humans are free? Partly because their logic and language preoccupation makes them say that either determinism or indeterminism is "true," and the other must be "false." This is the standard (but flawed) argument against free will.

But there is a deeper concern. If the origin of possibilities is truly random, have we lost the control needed to assert moral responsibility? Can the two-stage model provide a measure of control over the creative generation of alternative possibilities that does not make them pre-determined? Let us see.

The Standard Argument *Against* Free Will

Simple variations of this standard argument are found throughout the somewhat unsophisticated philosophical literature on free will,[18] and even in some of the most extensively cited work, for example, Galen Strawson's "Basic Argument on the Impossibility of Moral Responsibility."[19]

The standard argument has two parts.

If determinism is the case, the will is not free.

If indeterminism and real chance exist, our will would not be in our control, we could not be responsible for randomly caused actions.

The two-stage model provides the two essential requirements needed to defeat this standard argument

The first requirement is some indeterminism (objective chance) to break the causal chain of determinism and to generate creative thoughts and alternative possibilities for action. But this indeterminism must somehow not destroy our moral responsibility. It must not be the *direct cause* of action.

18 See Doyle, 2011, chapter 4, for dozens of examples.

19 Philosophical Studies: Vol. 75, No. 1/2, (Aug., 1994), pp. 5-24

Thus the second requirement is that our deliberations and evaluations are "adequately" (or statistically) determined, so that we can be responsible for our choices, so that they are "up to us."

"Adequate" (i.e., statistical) determinism means that the indeterministic alternative possibilities themselves are not the direct cause of our actions. The cause is the agent's decision.

Objective chance in the generation of alternatives means that at least some of the possibilities are not causally determined by immediately preceding events, so they are unpredictable by any agency, including us. They can then be the source of the creativity that adds new information to the universe.

Chance gives us the "free" in free will.

Adequate determinism gives us the "will" in free will.

Thoughts come to us freely. Actions go from us willfully.

We must admit indeterminism, but not permit it to produce random actions as some Determinists mistakenly fear.

We must also limit the determinism, but not eliminate it as some Libertarians mistakenly think is necessary.

The evaluation and careful deliberation of all the available possibilities, both ingrained habits and creative new ideas, can be recognized as "self-determination." This makes us the responsible "agent cause" of our actions.

But we must not thing that our "self-determination" was in any way pre-determined before we began to consider our possibilities. Self-determination is only "adequately and statistically" determined. It is not completely immune from random noise.

Compatibilists should be comfortable that the reasons, motives, feelings and desires of the agent are causal factors that were evaluated by the agent during the second-stage deliberations and the ultimate choice of an action.

This is all that is needed for the agent to accept what Robert Kane calls "ultimate responsibility" for the action.

But some event *acausality* is a prerequisite for any kind of *agent causality* that is not pre-determined by the moments before deliberations begin. This *acausality* is the quantum indeterminism, the

ontological chance, that accompanies the new information creation in the first stage of the two-stage model, where the agent freely generates alternative possibilities for action.

The two-stage model of free will proves that our actions are not predetermined, even from moments just before we begin thinking about freely generating new options for action.

We can summarize our criticism of the standard argument against free will in a few simple lines.

"Free Will" is really two independent stages that combine a limited indeterminism with a limited determinism.

First comes the "free" generation of alternative possibilities, then our adequately determined "willed" actions.

Our thoughts are free. Our actions are willed.

First "free," then "will."

Possible Worlds and Alternative Possibilities

In the twentieth century the study of modal logic (the truth conditions for statements about necessity and possibility) led to a model theory involving possible worlds. The philosopher David Lewis maintained there are an infinite number of possible worlds, all just as real for their inhabitants as our actual world. The physicist Hugh Everett III said that the world splits in two whenever a quantum experiment is performed.

Lewis and Everett were materialists and determinists. In their worlds everything is determined by the laws of nature and the fixed past. Each world has but one future. Free will is an illusion.

But Saul Kripke, who formulated the theory of possible world semantics for modal logic, described the use of possible worlds as representations of how our actual world might be. "'Possible worlds' are total 'ways the world might have been,'" he said, which means they can describe the *alternative possibilities* of our two-stage model for free will.

They are "counterfactual situations" in Kripke's sense, involving a single individual. Suppose the agent is considering five different courses of action. During the second stage of evaluation and deliberation only one of the five options (each a "possible world") will become *actualized*.

Note that Kripke's possible worlds are extremely close to one another, "nearby" in the sense of their total information content, the difference between them is very small amount of information compared to the typical examples given in possible worlds cases.

For typical cases of a free decision, the possible worlds require only small differences in the mind of a single person. Kripke argued against the thesis that mind and body (or brain) are identical. In this example, it would only be the thoughts in the mind of the agent that pick out the possible world that will be actualized.

Free Will and Creativity

Creativity requires that *new information* come into the world. It must be information that was not implicit in earlier states of the world. Information is only fixed in a deterministic universe.[20]

It is new information creation that explains *agent causality*.

When we create new information, we do it freely. Our thoughts are "free creations of the human mind," as Einstein said.

Humans are conspicuous creators and consumers of new information structures, altering the face of planet Earth. And we create the constructed ideal world of thought, of intellect, of spirit, including the invention of the laws of nature, followed by the discoveries that confirm them experimentally.

We are authors of our lives and co-creators of our natural world.

20 See appendix A.