Organon F 31 (1) 2024: 22–47
https://doi.org/10.31577/orgf.2024.31102
ISSN 2585-7150 (online)
ISSN 1335-0668 (print)

RESEARCH ARTICLE

Free Will: A *consensus gentium* Argument

William Hunt*

Received: 20 March 2023 / Revised: 1 November 2023 / Accepted: 7 January 2024

Abstract: This argument for free will is a probabilistic one based upon two conjectures: first, that of consensus; namely, that a large majority of people believe that they and others have free will and second, that *a priori* proofs against the existence of free will either fail or remain questionable. If these two conjectures hold, an inductive argument follows on the basis of beliefs founded upon justified auxiliary assumptions, assumptions that ensure a well-defined probabilistic relationship between the evidence of consensus and the proposition *free will exists* in an elaborated form. I will then demonstrate, through subjective Bayesian confirmation theory, that such evidence probabilistically confirms this proposition. Moreover, if one's prior degree of belief in the existence of free will is not very low - prior that is to consideration of the evidence - then, provided this evidence is factual, it is likely that one's resultant degree of belief in the veracity of the proposition is not only rational, but also compelling.

Keywords: Bayesianism; consensus; free will; libertarianism; probability.

* MF Norwegian School of Theology, Religion and Society
https://orcid.org/0000-0002-8206-3413
MF Norwegian School of Theology, Religion and Society, P.O.Box 5144 Majorstuen, 0302 Oslo, Norway
willghunt@gmail.com

1. Introduction

The free will debate is both ancient and voluminous and is fundamentally divided into three questions: what is free will, does it exist and, if it does, how can it be coherently explained? The objective of this paper it to address the second question – does free will exist? This question has been intensely debated with several conceptual refutations. In contrast, evidence for its existence has been proffered; however, there is an omission in this evidential deliberation. There has been no assessment of the probabilistic potency of this evidence on the likely truth of the proposition *free will exists*. Evidential arguments that circumvent this probability perspective can lead to an exaggerated view of the force of the evidence - I intend to address this lacuna. Thus, my primary objective, as the title suggests, is to provide a *consensus gentium* argument or *agreement of the people*: the majority. Herein, I extend its application beyond a simple majority to a probabilistic conception in terms of degrees of belief. Then, through Bayesian confirmation theory, I apply the evidence of consensus and other relevant facts to the free will proposition. This will demonstrate probabilistic confirmation of the proposition that establishes free will is more likely to exist than not.

Although it is likely that there would be a general consensus in support of free will, the application of Bayes' theorem provides objectivity to any probabilistic connection.

1.1. The free will proposition

To attempt to ascertain the probability that the proposition *free will exists* is true, a more encompassing proposition is required to give substance to its meaning. There are a number of interpretations of the meaning of free will and the following proposition is one that encompasses both the compatibilist and libertarian perspectives including my own:

> *h: Agents possess the capacity to make uncompelled reasoned choices between alternative possible actions so as to fulfil or resist a desire, whereby any resultant action or abstention to instantiate that choice is both intended and uncoerced.*

This free will proposition is commensurable with both the compatibilist and libertarian. For the compatibilist, the origin of the power that determines the agent's choice and any resultant action thereto is the causal nexus of a deterministic universe. By contrast, for the libertarian, the origin of the power that determines the agent's choice and any resultant action is the agent herself. For both parties, uncompelled choice and uncoerced action are necessary for predicating free will to the agent.

2. Probability

I now turn to the central theme of my argument – probability and the *consensus gentium* argument.

I contend that any argument for the existence of free will is primarily an evidential one, with the veracity of the free will proposition being subject to that evidence through a probabilistic analysis.

To ensure the proposition/evidence relationship is sound, any probabilistic analysis should be commensurable with the axioms of probability; herein, I apply the Kolmogorov axioms.

Probability theory includes a range of theories beyond the scope of this paper, but herein, I employ subjective probability.[1]

2.1 Subjective probability

Subjective probability is a form of epistemic probability which comprises two theories:[2] the logical theory and the subjective theory; (Gillies 2003, 37ff).[3] However, the logical theory is problematic as, although it complies with the axioms of probability, it relies upon the Principle of Indifference, a principle that leads to several paradoxes; (Gillies 2003, 33-49).

The subjective theory is based upon the personal credence someone gives to the chance of, in this case, a proposition being true. Warranted credence

[1] See Gillies (2003) for an analysis of the different theories of probability.

[2] Epistemic probability contrasts with objective probability of which there are two theories: frequency theory-Von Mises (1919) and propensity theory-Karl Popper (1959)

[3] Gillies also demonstrates that subjective probability is both necessary and sufficient for the axioms of probability; Gillies (2003, 59–64)

is expressed as a coherent degree of belief in a proposition given the evidence. Coherence is derived from the act of placing a bet, and defined in terms of avoiding a Dutch-book bet.[4] Like logical probability, credence is expressed as a numerical value between 0 and 1 on the probability continuum.

2.2 Conditional probability

Given a proposition h and evidence e, conditional probability is the probability of the truth of h given e expressed as Pr(h|e) and defined as:

$$\Pr(h|e) = \frac{\Pr(h \,\&\, e)}{\Pr(e)}$$

provided, $\Pr(e) \neq 0$

Herein, I employ conditional probability in the form of likelihoods, and for such likelihoods to be well defined, there are auxiliary assumptions that must be accounted for.[5] Auxiliary assumptions are crucial to this probabilistic analysis of free will and I consider them below.

2.3 Bayes' Theorem

A useful probabilistic tool in assessing whether some evidence provides justifiable credence in accepting a proposition to a certain degree is that of Bayes' theorem,[6] and the theorem is commensurable with subjective probability. The theorem is expressed as follows:

$$\Pr(h|e \,\&\, k) = \frac{\Pr(e|h \,\&\, k)\,\Pr(h|k)}{\Pr(e|k)}$$

---

[4] A Dutch-book bet is where odds are set by the bookmaker to win more money than the better, even if the better wins the bet. A Dutch-book is avoided by coherence with the axioms of probability. Coherence, as so defined, was proposed almost simultaneously by Frank Ramsey,1926 and Bruno De Finette,1930

[5] A likelihood is a conditional probability function of the form Pr(e|h & k); where, in this case, the probability of the evidence e is conditional on the assumed truth of the proposition h and background knowledge k, which include auxiliary assumptions that create a well-defined probabilistic relationship between h and e.

[6] For challenges to the theorem and their defence see Earman (1996, Ch.4)

and can be reformulated, in its comparative form; i.e., h compared to ¬h (mutually exclusive and exhaustive propositions), using a likelihood ratio as follows:

$$\Pr(h \mid e \,\&\, k) = \frac{\lambda \times \Pr(h|k)}{(\lambda \times \Pr(h|k) + (1 - \Pr(h|k)}$$

Where the likelihood ratio λ is:

$$\lambda = \frac{\Pr\,(e|h \,\&\, k)}{\Pr\,(e|\neg h \,\&\, k)}$$

The function k represents background knowledge that also includes the all-important auxiliary assumptions.

This reformulation of Bayes' theorem is useful when used with subjective probability as likelihood ratios are much easier to assess subjectively than individual likelihood values.

Pr(h|e & k) is the posterior degree of belief in the proposition h; that is, the new degree of belief that would be formed if the person *conditionalised* on the evidence e with respect to h (see Bayesian conditionalisation below). k is background knowledge, which includes the auxiliary assumptions.

It can be seen that if λ is greater than 1, then probabilistic confirmation follows; i.e. Pr(h|e & k) > Pr(h|k). If it is less than 1 then probabilistic disconfirmation follows; i.e. Pr(h|e & k) < Pr(h|k). If it is 1 then there is no confirmation or disconfirmation; i.e. Pr(h|e & k) = Pr(h|k).

3. A *consensus gentium* argument

Having outlined my probabilistic methodology, I now turn to the probabilistic analysis for the existence of free will – a *consensus gentium* argument. I will argue for, and present values for the functions in the likelihood ratio form of Bayes' theorem above, with particular attention to the evidential function e – this represents the consensus element of the *consensus gentium* argument. In addition to this evidential element, I will focus on the auxiliary assumptions essential for any conditional probability analysis of this sort; that is, ensuring there is a well-defined probabilistic relationship between h and e.

3.1. Propositions h, ¬h and e

The free will proposition h is more than just *free will exists*, as it can imply different interpretations with different existential possibilities and only needs one case in an infinite universe to confirm its truth, which does not capture its intended meaning. Thus, for this Bayesian argument I expose h, the aforementioned free will proposition, and its negation ¬h, to the evidence of consensus:

> h: Agents possess the capacity to make uncompelled reasoned choices between alternative possible actions so as to fulfil or resist a desire, whereby any resultant action or abstention to instantiate that choice is both intended and uncoerced.
>
> ¬h: Agents DO NOT possess the capacity to make uncompelled reasoned choices between alternative possible actions so as to fulfil or resist a desire, whereby any resultant action or abstention to instantiate that choice is both intended and uncoerced.

A *consensus gentium* argument would normally be considered fallacious; evidence of consensus is, *prima facie*, subject to prejudice and can be unreliable. However, herein I justify its application by employing probability theory with robust auxiliary assumptions.

The vast majority of participants of a general survey on free will would be unaware of the nuances of the free will debate to make an informed decision. In fact, a loss of precise conceptual correspondence between individuals is likely to lead to confusion and imprecision, making any data untrustworthy. Thus, the population of this *consensus gentium* argument should comprise a body of participants likely acquainted with the ebb and flow of the free will debate. Given the above, e is defined as:

> e $=_{df}$ *The mean degree of belief in free will expressed as a probability quotient within a given population acquainted with the free will debate is greater than 0.5 – more likely than not.*[7]

[7] I should add that e represents a consensus as independent agreement, not consensus by cooperative agreement as with intersubjective probability.

3.2. Auxiliary assumptions

Subjective Bayesianism is based upon subjective probability which, in this analysis, is a degree of belief in a proposition based upon evidence – a truth-conducive interconnection enabled by the auxiliary assumptions.[8]

Adopted auxiliary assumptions have themselves to be justifiable as there is a danger of bulking-up the likelihood with unjustifiable auxiliary assumptions to leverage the probability in one's favour. "What is needed is not the invention of auxiliary propositions…but the identification of auxiliary information that is independently supported." (Sober 2008, 168) I provide five auxiliary assumptions as follows:

3.2.1. Naturalistic realism

Naturalistic realism assumes that, given all the possible explanations of reality, the best explanation is that provided by current scientific theory, which can change over time – it is defeasible; (Kuhn 2012). It projects its conception of reality beyond that which is observable, but is still commensurable with observable evidence. The science of cognitive psychology accepts the reality of the mind, in particular intentional agency, and is a form of naturalistic realism. Moreover, it is a widely accepted evidence-based science consistent with that of other human and social sciences.

The tenets of cognitive psychology (particularly intentional agency) under the lens of naturalistic realism would be a substantial auxiliary assumption in the likelihood and prior functions Pr(e|h & k) and Pr(h|k).

3.2.2. Evolution

The emergence of consciousness, instincts and cognition in early life on Earth provided such life with improved survival chances. As the evolutionary course progressed, instincts and beliefs emerged that interfaced with

[8] The auxiliary assumption requirement is associated with the Duhem–Quine thesis. The thesis is a combination of Pierre Duhem's 1904/5 thesis and Willard Van Orman Quine's 1951 article Two Dogmas of Empiricism. In short, it is impossible to test a scientific hypothesis in isolation. It requires several background assumptions termed auxiliary assumptions or bundles of hypotheses; see Gillies (1993, 98-116)

environmental dangers, food supply, predators and many other advantages and threats, providing further survival chances and subsequent adaptation. Moreover, if those beliefs were true rather than false, survival chances would improve even further. Thus, the cognitive ability and motivation to harbour true beliefs delivers selective pressure on both animals and early humans which would have manifested itself in finding shelter, socialization, communication and so on. Indeed, Anabela Pinto argues that the complex beliefs that modern humans hold share a relationship with animal beliefs indicating biological roots of belief formation by adaptation. This, she argues, points the way to an evolutionary explanation for our complex linguistic belief concepts; (Pinto, 2022, 22).

Clearly, all our beliefs are not true, but we are motivated to harbour true rather than false beliefs and change them if shown to be false – an echo from our distant biological roots. Moreover, in modern life there is personal developmental pressure from a whole range of sources to form and harbour true beliefs over false ones. These two doxastic factors help define a probabilistic relationship between e and h in the likelihood Pr(e|h & k). This is because a belief in free will assumes the truth of h rather than ¬h, and true beliefs are more likely with the evolutionary and developmental influences than without.

3.2.3. Phenomenology

The phenomenal experience of free will in terms of the first-person sense of being in control, losing control and regaining control (control-phenomena) are very common experiences for us all, albeit sensed differently. Similar experiences provide a sense of temporal and spatial awareness, self-awareness, social awareness and a host of other essential senses.

Control-phenomena suggests an ontological as well as qualia-logical content – it provides a lens on the power of self-determination possessed by us. This is central to the social science of cognitive psychology as outlined above. However, do control-phenomena ensure the existence of self-determination in the same way that self-awareness ensures our existence?

Self-determination implies that we can control our decision making. Loss of self-control means we are under the spell of our emotions, and decision-making becomes less rational – we're not fully in control of our decision making. If we're not fully in control of our decision-making then we cannot

be said to be self-determining our decisions – our emotions have taken over for example. However, if we begin to regain self-control, self-determination progressively returns. Even though self-control does not entail self-determination, the two concepts are related and control phenomena indicative of the reality of self-determination.

We have a distinct conscious awareness of losing control when we are at the mercy of causal-power such as the emotions of anger and fear, and we sense its reinstatement with the shift back to self-control - all frequently manifested in behaviour with concomitant third-party interpretation. Indeed, first-person experiences of pain, control, irritation, affection etc are frequently manifested in behaviour, and introspective predicates develop from that correlation; (Wittgenstein 1976, sec.244).

Token control-phenomena can be subject to confabulation in terms of scaling, particularly with retrospective rather than concurrent recall. Moreover, cognitive masking during loss of control such as shouting, or being spoken to whilst regaining control can weaken recall of the control-phenomena. Notwithstanding, this type of introspective experience is similar to pain in terms of recalling intensity - both types of experience are incorrigible, even if detailed recall is confabulated; (Shanahan 2010, 67-89).

The auxiliary assumption of control-phenomena helps create a well-defined probabilistic relationship with the function $Pr(e|h \,\&\, k)$ – also with the assessment of the prior $Pr(h|k)$.

3.2.4. Blame and liability

When our choices are instantiated by us they can lead to blame and legal liability if there is a breach of moral rules or law respectively. The rule of law is a global form of social control even though laws can vary from country to country, but all are predicated on the understanding that if members of a social group breach such laws they will be held responsible. Moreover, many societies hold to the maxim that being responsible for breaking the law is a necessary criterion for blame and punishment, and being responsible means that actions are down-to-them, they are the source of the breach that could have been avoided by an alternative choice of action.

Although it could be argued that law is solely a means of social control, and punishment merely deterrence, the reality is that, in addition to any

social control, blame and punishment is fundamentally retributive. Indeed, Daniel McDermott argues that, analogous to financial debt, criminals incur a non-consequentialist "backward looking" moral debt to their victims (including society in some cases) and proportionate punishment, as retribution, represents a settlement of that debt; (McDermott 2002, 439–464).

Blame and punishment as retribution presupposes liability that, in turn, presupposes responsibility which then implies personal control in our choices of action unless proved otherwise; (Pereboom 2014, 153–160). It is the global proliferation of moral rules and the rule of law, together with assumed self-determination in breaches of them that, as an auxiliary assumption, helps create a well-defined probabilistic relationship between h and e in the likelihood Pr(e|h & k). Moreover, this auxiliary assumption has application to the prior Pr(h|k) as background knowledge k would include knowledge of moral rules and the rule of law and the related assumption of intentional agency.

3.2.5. Scholarly error

There can be widely held false beliefs among scientists; for example, the Newtonian concept of gravity or the hypotheses of phlogiston, vitalism and luminiferous ether. The cause of these false believes was primarily an absence of relevant facts, faulty or limited equipment and a lack of insight rather than self-delusion or mental weakness. Nevertheless, historical precedent and peer/societal pressure could have contributed, and the same could apply to free will; (Kuhn 2012, 66–76). Neuroscience or possibly physics may eventually fill any evidential and explanatory lacunae and refute the existence of free will – it is a defeasible notion.[9] In fact, defeasibility is an assumption of naturalistic realism and, given the history of scholarly error, an auxiliary assumption arises providing probabilistic definition to the function Pr(e|¬h & k). Residing in background knowledge k, this assumption may also affect the assessment of the prior Pr(h|k).

[9] The time-lapse in the Benjamin Libet et al (1983) experiment came close to such falsification with reflexive reactions; however, not with reflective reactions; also see List (2019, 141-147) and Rolls (2012)

4. Valuation

I shall now apply the presuppositions of subjective probability to ascribing values to the conditional probability functions of the likelihood ratio form of Bayes' theorem.

4.1. Pr(h|k): Prior belief in h

Because there are differences of opinion concerning the truth of the free will proposition, Pr(h|k) could be allocated 0.5 on the basis of the Principle of Indifference of logical probability. However, because of the paradoxes associated with the Principle, in this case the Book Paradox,[10] I resort to a subjective evens betting position and still set Pr(h|k) = 0.5. I do this on an assumption of no prior knowledge of the auxiliary assumptions.[11]

4.2. Pr(e|h & k)/Pr(e|¬h & k): the likelihood ratio λ

Given the substantiality of the auxiliary assumptions that provide probabilistic definition between h and e, I would argue that Pr(e|h & k) >> Pr(e|¬h & k) and as such λ would be greater than 1 leading to confirmation; i.e. Pr(h|e & k) > Pr(h|k). Others may prefer Pr(e|h & k) > Pr(e|¬h & k) or Pr(e|h & k) >>> Pr(e|¬h & k).

Despite the validity of subjective probability, providing a precise value of λ would be somewhat arbitrary. Notwithstanding, given the subjective assessment Pr(e|h & k) >> Pr(e|¬h & k), my evaluation of ratio λ is not less than 1.5 but certainly no greater than 3; i.e. a range of possible values.

Given Pr(h|k) = 0.5 (on the basis of a subjective *evens* bet referred to above) and λ = f(y) which lies in the range [1.5, 3] then, from the comparative form of Bayes' theorem above:

---

[10] The falsity of any one of the criteria of h would ensure ¬h. Thus, there could be range of ¬h definitions, each with a probability of 0.5 given the Principle of Indifference which would breach axiom 1; see Gillies (2003, 37f)

[11] The problem of old-knowledge would challenge the use of Bayes' theorem in this application; see Glymour (1980, 86). However, see Howson & Urbach (1991, 270f) for a defence.

$$\frac{y \times 0.5}{(y \times 0.5)+(1-0.5)} = \Pr(h \mid e \,\&\, k) \text{ which lies in the range } [0.60, 0.75]$$

Thus, based upon an assumption of the evidence of consensus, the proposition h *free will exists* as so expressed is more likely true than false with a subjective probability value of between 0.60 and 0.75 – probabilistic confirmation. However, for Bayesian-conditionalisation to occur (coming to believe the above result) the evidence must be actual rather than assumed.[12]

4.3. Bayesian Conditionalisation

It is the reality of e that gives the Bayesian reason to accept the posterior probability value and conditionalise on that value, and it is to surveys I turn to provide that evidence.

There are two surveys that plainly fulfil the *acquainted with the free will debate* criterion within e: The *2009* and the *2020 Philpapers surveys*, both conducted by David Bourget and David Chalmers who posed questions to target audiences comprising philosophers on a range of philosophical issues, including one on free will (Bourdet and Chalmers 2013; 2020; 2023). A brief outline of the surveys is as follows:

(i) *Target population (2009)*: The survey was taken by 3226 respondents with philosophical backgrounds from Australasia, Canada, Europe, UK and US.
Target population (2020): The survey was taken by 7685 respondents with philosophical backgrounds from New Zealand, Canada, Ireland, UK and US.

(ii) *The free will question (2009)*: Accept or lean towards: compatibilism, libertarianism, or no free will?
The free will question (2020): Accept or lean towards: compatibilism, no free will, or libertarianism?

(iii) *Numbers answering the free will question (2009)*: 931
Numbers answering the free will question (2020): 1758

[12] For a discussion on Bayesian conditionalisation see Howson and Urbach (1991, 67f).

(iv) *% Results (2009)*: compatibilism - lean towards: 226 & accept: 324; libertarianism – lean towards: 56 & accept: 72; no free will – lean towards: 62 & accept: 53; and 'other' 138 of which 38 are only relevant to the question - being the agnostic response.

(v) *% Results (2020)*: compatibilism – lean towards: 490 & accept: 550; libertarianism – lean towards: 138 & accept: 193; no free will – lean towards: 102 & accept: 95; and 'other' 190. The 'other' responses are more varied than the 2009 survey, but the only clear applicable result being the agnostic one of 80 responders.

What is interesting with both these surveys, is the number of participants that did not answer the free will question (2009: 2295; 2020: 5927) indicating agnosticism – a degree of believe of 0.5 on the probability continuum. In terms of degrees of belief, these agnostics cannot be ignored and neither can the disbelievers nor the 'other' group. Indeed, 38 responders of the 'other' group in the 2009 survey were agnostic; there is no indication that the remaining responses were relevant to the question, and therefore should not be included in the analysis.

What is also interesting from these surveys with regard to degrees of belief, is the division between 'lean towards' and 'accept'. Thus, modelling the probability continuum into equal subintervals to represent 'lean towards' and 'accept' for disbelief and belief, with agnosticism at the midpoint of the continuum we have: 0…0.17…0.34…0.50…0.67…0.84…1.[13]

Applying the above model to the 2009 survey, a mean degree of belief of ≈ 0.53 results and applying the above model to the 2020 survey, a mean degree of belief of ≈ 0.53 results. These two results are virtually identical adding further credence to warranted Bayesian conditionalisation of Pr(h|e & k) whose value lies in the range [0.60,0.75].

The individual degrees of belief among the respondents will vary despite being within the accept, lean towards or agnostic groups. Consequently, the above probability continuum model may not be precise; notwithstanding,

[13] The probability calculus assumes that probabilities are real numbers and each probability on the continuum should, theoretically, be represented by an infinite decimal (e.g. 0.1 is given by 0.999…∞) because the probability space between any one point on the continuum and another is infinitely divisible. However, with subjective probability the calculus is an approximation as values are vaguer.

the values selected in the model are an even distribution of the continuum that reflect the differing degree of belief modes. Even skewing the model towards disbelief, say 0..0.1..0.17…0.50..0.60..0.67…1 yields a mean degree of belief greater than 0.5 for both the 2009 and 2020 surveys.

The above mean degree of beliefs results for the 2009 and 2020 surveys summate the results of the compatibilists and libertarians. However, this may not be justified given the differences between the two camps. Supporting facts in Bayesian conditionalizing are not intended to be entailments, but are persuasive facts to a greater or lesser extent. Thus, it is a question of whether h is congruent with both compatibilist and libertarian views of free will. I believe it is, and therefore contend that it is sound to combine both the compatibilist and libertarian results in the surveys, and that the two surveys provide justifiable evidence to Bayesian conditionalise on Pr(h|e & k) whose value lies in the range [0.60,0.75].

There are other free will surveys, but none as convincing and targeted as the two above.[14]

5. Preliminary conclusion

My preliminary conclusion is that, based upon the evidence of consensus as so defined and the justification of the auxiliary assumptions, the posterior value of the proposition *free will exists* as so expressed is greater than 0.5 – free will is more likely to exist than not. This does not ensure the truth of the proposition, and there is room for evidential refutations that neuroscience may provide. However, there are *a priori* refutations that threaten this preliminary conclusion.

[14] See for example Wisniewski et al (2019), where their survey found an 82.33% belief in free will in the US and 85.44% in Singapore. Also, in January 2015 Gary Stix carried out a survey for Scientific American with 4672 responders from the US including some from France, Australia, New Zealand, Kuwait, Israel, the Philippines and India. 59% believed in free will and 41% disbelieved.

6. A libertarian interpretation

There are different interpretations of free will that are commensurable with the free will proposition h, mine is best described as non-causal libertarianism. This is the perspective I employ in addressing the *a priori* refutations of free will below; as such, a more detailed explanation is required.

By non-causal libertarianism I assert that persons possess a distinct power of self-determination compatible with causal determinism should it exist –a compatibilist interpretation of free will.

This interpretation implies two freedoms, *the freedom to choose otherwise* and *the freedom to do otherwise*; that is, free will and free action – interrelated and goal directed concepts. Indeed, this power to have *chosen otherwise* is central to a libertarian perspective – a *multi-way* power; (Pink 2019a, 268). Indeed, we sense this multi-wayness when reflecting upon alternative possibilities at the point of choice. In addition to the freedom to choose otherwise, the possibility for voluntary action to fulfil a desire, including abstention is also fundamental to libertarianism.

Intentional agency is another key criterion to the libertarian perspective despite its superfluous presence for the epiphenomenalist, incompatibilist or even the classical compatibilist.[15] Intentional agency is a goal directed choice followed by a goal directed voluntary act; (Pink 2019a, 259-266).

The most overt threat to the libertarian is the hypothetical problem of causal determinism. However, there is an alternative to the power of causal determinism; that is, the power of self-determination - an intrinsic power that I term *will-power* as contrasted to *causal-power*.[16]

Will-power is a 'difference making' intrinsic power that is,[17] like causal-power, difficult to define in an ontic sense, but unlike causal-power it is not

[15] See Hobbes (1841, XX) - although Hobbes eschewed the will as the cause of voluntary action, he viewed free will as simply the unimpeded satisfaction of desires.

[16] I use the term *will-power* only as a contrast to *causal-power* not in the usual sense of fortitude.

[17] By 'difference making' or "matterers" as Helen Steward terms it, I refer to facts that make a difference to an outcome as contrasted with effects from dynamic causal forces; Steward (2014, 212ff). Christian List also adopts this notion; List (2019, 131–140). Both Steward and List apply the notion to causation, herein I apply it to will-power.

realized by observing regularities in nature as Hume would have it; (Hume [1739] 1985, III, 117-123). In fact, our habits can be regular, but our choices are frequently not. Indeed, as well as lacking such regularities the difference between will-power and causal-power is stark - causal-power excludes goal-directed intention, choice and multi-wayness, blame and moral responsibility. These differences between will-power and causal-power remain even if the concept of causation is expanded from its dynamic character such as *the wind blew the chimney off*, to include making a difference such as *inflation soared because wages increased*. Compare this to *I refused a drink because I'm driving* – the above differences still apply.

From observing causal-power in the natural world, there is a temptation to predicate libertarian free will as causal - *agent-causal libertarianism.*[18] However, as I argue above, will-power is so different to causal-power that it warrants its own designation rather than being a sub-category of causal-power.

This power-difference perspective assumes that both free will and causation exist, with each having powers to bring about change in different ways. However, although related in this sense, the two powers cannot be conflated. Thus, the tag *non-causal libertarianism* has application to this power-difference perspective.

6.1. The mechanics of libertarianism

Libertarianism is intuitively compelling given our everyday phenomenological experiences, and some evidence does suggest a relationship between specific conscious decision making and concomitant action.

6.2. Correlation

The essence of the libertarian perspective is that this conscious decision making has a power over and above the causal nexus in which the neural networks are seated. The threat to libertarianism is that such mental states are superfluous to the train of the causal nexus and that there is only a correlation between causal neural activity and conscious decision making.

[18] See Pink (2019a; Ch.14) for a critique of agent-causal libertarianism.

This correlation perspective has gained credence in neuroscience from the notion of the neural correlate of consciousness (NCC) pioneered with the use of fMRI scanning together with reported conscious experiences. (Charmers, 2000, 17-39) However, libertarians need more than a mere correlation, they need an instrumental power that emanates from the agent.

6.3. Integrated information theory

An alternative explanation to correlation is that given by integrated information theory (IIT); (Tononi, 2004), derived and explained by a set of five axioms and resultant postulates;[19] (Tononi et al, 2023, 3-5). With IIT, the information element relates to neural systems functioning to reduce experiential uncertainty by ruling out experiences from a range of possible ones – *differentiation* or a not this or that scenario; (Seth, 2021, 52f).

The integration process is a function of the neural system as a self-causal unified whole rather than isolated individual systems; i.e. parts of the system affect other parts and, in turn, are affected by them – a cause/effect interdependence. There is synergy with such integration; i.e. extra information – Φ being the measure of this holistic extra; at least in principle. When such integration reaches a high level (maximally irreducible conceptual structure (MICS))[20] the system is conscious – a self-generated emergent property of integrated information.

With IIT, there is an identity assertion - consciousness is MICS,[21] and has variable Φ depending of the level of integration; dreaming having low Φ. MICS is self-generated, intrinsic to the neural system giving it potential for a libertarian explanation for free will; in short, it is the power to make a difference to itself. Indeed, MICS supervenes on the substrate of consciousness; i.e. the integrating neural system. In this way, MICS should be seen as a holistic state comprising the phenomenal and neurological. Indeed, IIT claims just that, as consciousness is deemed identical to MICS; i.e. a maximal integrated system. Thus, consciousness is inextricably integrated

[19] For Tononi, the move from axiom to postulate in not deductive, but abductive; (Tononi et al, 2023, 3)

[20] Irreducible in the sense of greater than the sum of its parts.

[21] In the sense of numerical identity, despite a difference in meaning.

with the neural system. As such, MICS possesses the power to change the system where the phenomenal alone could not.

IIT provides a compelling explanation of how an agent possesses intrinsic self-generated power to make and fulfil her choices – libertarian free will.

IIT is a hypothesis, and as such, requires empirical evidence in support, and given that MICS is essentially intrinsic to the person, observation, other than measuring the firing and integrating of neural systems, is private. However, comparative brain functions provide persuasive evidence for IIT. Despite the cerebellum possessing c.70 billion neurons it has been shown to have no contribution to the conscious state. In contrast, the thalamocortical system possesses c.16 billion neurons yet is the core of consciousness. The fundamental difference between the two brain systems is that the thalamocortical system has a complex nexus of interconnecting neural fibres compared to the cerebellum, enabling a very high level of integration – supporting evidence of IIT; (Tononi, 2004, 10ff).

Having espoused my non-causal libertarian perspective of free will, I turn my attention to the *a priori* refutations of free will and my replies thereto.

7. Refutations and Replies

There are a number of objections, with concomitant defences, to the existence of free will and the veracity of subjective Bayesianism, and I have referred to some in this paper. However, there are potential *a priori* refutations of the free will proposition and Bayes' theorem has no application where the probability of the proposition is 0 (certainly false), and these *a priori* arguments threaten just that. Similarly, an ontology that views only evidence as real and relegates self-determinism and intentional agency to only useful but unreal constructs ensures the failure of a Bayesian analysis of free will.

Currently, there are no *a priori* arguments that establish the certain truth of the free will proposition. However, there are arguments that suggest the free will proposition is certainly false, or unreal in the case of scientific instrumentalism. I examine such arguments as follows:

7.1. Strawson's infinite regress argument

Galen Strawson's argument assumes a premise that free will entails ultimate moral responsibility for acts freely instantiated; (Strawson 1986, 292ff).[22] This implies that the agent's intentional choice is, itself, freely preferred for certain reasons. In this way moral responsibility for the ensuing act can be predicated to the agent. However, these reasons must be persuasive enough for the agent to prefer her choice, i.e. "principles of choice…preferences, values, pro-attitudes, ideals…" (Strawson, 1986, 25) Nevertheless, for moral responsibility to hold, such persuasiveness cannot just emerge; it too must have principles of choice that provide reasons persuasive enough for the agent to be persuaded that her choice is best. Thus, to ensure moral responsibility and hence the agent's free will, an infinite regress of reasoned persuasion arises along with recursive moral responsibility which is impossible. Consequently, moral responsibility and therefore free will is forfeit; (Strawson 1986, 24f).

Our desires, wants and passions are not reasoned into existence, but just come over us as a result of our instincts, observations and the ever developing autobiographical-self; (Søvik 2018,106-126). The agent becomes motivated to satisfy those desires, wants and passions, and intentional agency aims to do just that. There is no prior persuasion that requires justification for those desiderata; they just come over her. Her goal directed voluntary action to satisfy the desiderata follows. It is at that point that moral responsibility has application, not for any self-reflective mental processes that rationalize the desiderata into mental existence - the locus of responsibility is the act or abstention.[23]

7.2. Incompatibilist refutations

Although, there are arguments that deny determinism on the grounds that there are no universal laws of nature; (Cartwright 1999), there are incompatibilist arguments against the existence of free will based upon the truth of determinism.

---

[22] Ultimate responsibility implies the agent is the sole originator of the act.

[23] See Mele (1995, 223-225) for a counterfactual example.

Established incompatibilist arguments are the Consequence Argument, the Origination Argument and the Mind Argument. The Consequence Argument, championed by Peter van Inwagen, claims that the invariable state of past events (fixity of the past) together with the laws of nature, determine all future events; (Van Inwagen 1983, 16). Consequently, volitional and voluntary acts are illusionary. He structured this argument formally, and there are challenges to the validity of his argument; (Van Inwagen 2002, 158–177).[24]

The Origination Argument is also based on the truth of determinism and states that if determinism is true, an agent's volitions do not originate with her but are extrinsically caused – the agent is not the ultimate source of her volitions. Given this, and assuming the necessity of origination for free will, determinism is a serious defeater; (Kane 1996,79f).

The mind argument is so called because it challenges the existence of free will by the effects of determinism (and equally indeterminism) on the mental-state of freely choosing. The argument against free will has three strands; the first strand echoes the Consequence Argument, claiming that our choices are outside of our control and pre-caused by the progress of the causal nexus through time. The alternative is indeterminism, and if this means a random setting, then free will is still deniable – apparent actions are really just chance events. The second strand claims that volitional acts and voluntary acts are not acts at all unless they have a prior agential cause. The third strand concerns the action of choosing between alternative possibilities. Again, this echoes the Consequence argument as, given determinism, alternative possibilities are not possible; (Van Inwagen 1983, 126-152).

If true, self-determination as a *sui generis* power defeats all three arguments from incompatibilism. Determinism is more precisely termed *causal determinism*, and will-power, with is distinctive properties, is not causal and therefore, is not subject to the power of causal determinism. Thus, if determinism is true as a universal causal system, the agent can still intervene in its component causal sub-systems by her voluntary acts; (Ismael (2016, Chs. 4 & 5).

[24] Also see the Agglomeration Argument; McKay and Johnson (1996, 113–122)

Motivated by passive desires, wants and passions and influenced by the autobiographical self, the agent's will-power originates from her. Hence, the causal-nexus does not determine the agent's choices – she does. Further, voluntary acts are acts, not because of agential causation, but because of agential self-determination and, as I argue above, multi-wayness is a characteristic of will-power unlike causal-power. Thus, alternative possibilities are feasible irrespective of the truth of determinism. Also, indeterminism as viewed as chance events, is not a challenge to free will, as the power of self-determination eclipses any associated random possibilities.

There are additional replies to the three arguments from incompatibilism other than the power of self-determination. Indeed, incompatibilism and replies to it are a central item on the free will agenda but beyond the scope of this paper.

7.3. Instrumentalism

A scientific perspective in contrast to naturalistic realism is instrumentalism which is also evidence based, but it does not assert a true unobserved reality beyond the specific evidence. Any inference to the best explanation upon such evidence is based upon utility - the more useful the inference the more worthy its adoption. Thus, the instrumentalist would likely argue that cognitive psychology explains behavioural evidence, but such evidence is all that is real. Extrapolation from that evidence to intentional agency is not an assertion of reality, but an assertion of instrumentality in that it is practically useful; (List 2019, 74–77).

Prima facie it appears that instrumentalism, if sound, is an auxiliary assumption supporting the likelihood $Pr(e|\neg h \,\&\, k)$ as free will is a hypothesis based upon evidence. The hypothesis can be useful but cannot, itself, be considered true and in this likelihood the hypothesis is not considered true. However, instrumentalism does not consider h and ¬h in ontological terms, only in terms of their usefulness as constructs - the only component in the above Bayesian analysis that is real is e. So if instrumentalism is considered a sound scientific perspective of reality, then existential predicates do not apply to h and ¬h and the posterior function $Pr(h|e \,\&\, k)$ is otiose. However with instrumentalism, there is an overweighted emphasis

on distinct observation over the clear theoretical implications of evidence, bringing a vagueness to defining reality.

In summary, if instrumentalism is adopted, then explanatory power and correct predictions based upon non-existent scientific constructs seems to be more miraculous than rational; (Putnam 1975, 72f). Indeed, "Experimental physics provides the strongest evidence for scientific realism. Entities that in principle cannot be observed are regularly manipulated to produce new phenomena and to investigate other aspects of nature." (Hacking, 1982, 71)

The notion of self-determination as will-power casts doubt on the soundness of the potential *a priori* arguments against free will, and instrumentalism is a questionable scientific principle. Based on the above replies, I do not believe that these refutations ensure that the probability of the free will proposition is zero, or in the case of instrumentalism, that Bayes' theorem has no application.

8. Conclusion

The objective of this paper was to substantiate my contention that the debate concerning the existence of free will is essentially an evidential one rather than conceptual one; thus, lending itself to a probabilistic analysis. From an introduction to subjective probability, I moved my analysis to subjective Bayesianism applied to the free will proposition based upon the evidence of consensus – a *consensus gentium* argument.

I argue that evidence from consensus is only sound if the population selected is cognizant of the issues relevant to the question posed to it, in this case, *does free will exist*?, and the free will debate is both extensive and complex. Moreover, consensus itself is a belief orientated notion, and belief comes in degrees. Thus, I adjusted the evidential proposition in Bayes' theorem to account for these two nuances. I also employed a likelihood ratio in the comparative form of Bayes' theorem.

To ensure a well-defined probabilistic relationship between the free will proposition and its negation and the evidence of consensus, I introduced a set of auxiliary assumptions. These assumptions also had application to the prior belief in the free will proposition given background knowledge. Thus,

I examined five auxiliary assumptions: naturalistic realism, evolution, phenomenology, blame/liability and scholarly error. From Bayes' theorem, together with these auxiliary assumptions, I concluded that the posterior probability value of the free will proposition on the evidence would fall within the range [0.60,0.75]. However, for this result to be credible, such a theoretical analysis needs evidential support itself – Bayesian conditionalisation. Thus, I introduced the results of two surveys that supported this posterior value and concluded that the posterior probability value of the free will proposition on the evidence of consensus would fall within the range [0.60,0.75], so it is more likely than not that free will exists.

There are evidential challenges to the free will proposition that I referred to throughout the paper. However, a more serious challenge to the proposition loomed – *a priori* refutations. To reply to these refutations, I first provided my interpretation of free will that is congruent with the free will proposition h - *non-causal libertarianism*. This interpretation indicates a personal power possessed by free agents to make choices between alternative possibilities and to instantiate those choices. A power manifesting properties different to that of causation - I term this power *will-power* in contrast to *causal-power* – interrelated concepts that cannot be conflated. From this differentiation, it was clear that my argument would take on a compatibilist perspective, at odds with incompatibilism and event/agent-causal libertarianism. Given the threat of incompatibilism, I provided a brief insight into the likely neural mechanics of libertarianism – integrated information theory (IIT).

The application of Bayes' theorem presupposes that the probability of the proposition is not zero, and there are several *a priori* refutations of the free will proposition. I considered the primary ones and applied the concept of libertarian self-determination in defence of them, arguing that there is doubt as to their *a priori* status. I also cast doubt on the veracity of an instrumentalist perspective of reality that would have excluded a Bayesian analysis of free will.

Although there is strong evidence in support of the free will proposition, it does not ensure its truth. In fact, although my analysis demonstrates that the probability of the truth of the free will proposition is fairly high, this also implies there is a probability that it is false. The proposition is

defeasible, and new evidence, say from neuroscience and/or quantum science, may change this true/false-balance either way.

References

Bourdet, David. and David J. Chalmers. 2013. "What do Philosophers Believe?" *Philpapers*: https://philpapers.org/archive/BOUWDP

Bourdet, David and David J. Chalmers. 2020. "Philosophers on Philosophy." *The 2020 PhilPapers Survey*: https://survey2020.philpeople.org/survey/results/all

Bourdet, David. and David J. Chalmers. 2023. "Philosophers on Philosophy." *The 2020 PhilPapers Survey* Philosophers' Imprint 0: 0: 1-53 BOUPOP-3 (philarchive.org)

Cartwright, Nancy. 1999. *The Dappled World*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139167093

David, Chalmers. 2000. *What is a Neural Correlate of Consciousness*? In *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, edited by Thomas Metzinger, 17-39. Cambridge, Massachusetts: The MIT Press.

Earman, John. 1996. *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*. Cambridge, Massachusetts: MIT Press, first published 1992.

Gillies, Donald. 1993. *Philosophy of Science in the Twentieth Century: Four Central Themes*. Oxford: Blackwell.

Gillies, Donald. 2003. *Philosophical Theories of Probability*. London: Routledge.

Glymour, Clark. 1980. *Theory and Evidence*. Princeton: Princeton University Press.

Hacking, Ian. 1982. "Experimentation and Scientific Realism." *Philosophical Topics*, 13(1): 71-87. http://dx.doi.org/10.5840/philtopics19821314

Hobbes, Thomas. 1841. *The Questions Concerning Liberty, Necessity and Chance*, In *The English Works of Thomas Hobbes*, vol. 5, edited by Sir William Molesworth, Bart. London: John Bohn, first published 1654-1656.

Howson, Colin and Peter Urbach. 1991. *Scientific Reasoning: The Bayesian Approach*. Illinois: Open Court Publishing Company, first published 1989.

Hume, David. 1985. *A Treatise of Human Nature*. London: Penguin Classics, first published 1739.

Inwagen, Peter van. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.

Inwagen, Peter van. 2002. *Free Will Remains a Mystery*, In *The Oxford Handbook of Free Will*, edited by Robert Kane, 158-177. New York: Oxford University Press.

Ismael, Jenann. 2020. *How Physics Makes Us Free*. Oxford: Oxford University Press, first published 2016.

Kuhn, Thomas S. 2012. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, first published 1962.

Libet, Benjamin et al. 1983. "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness Potential). The Unconscious Initiation of a Freely Voluntary Act." *Brain*, 106: 623–642. http://dx.doi.org/10.1093/brain/106.3.623

List, Christian. 2019. *Why Free Will is Real*. Cambridge, Massachusetts: Harvard University Press.

McDermott, Daniel. 2002. "Debts to Society." *The Journal of Political Philosophy* 10(4): 439-464. http://dx.doi.org/10.1111/1467-9760.00160

McKay, Thomas and David Johnson. 1996. "A Reconsideration of an Argument against Compatibilism." *Philosophical Topics* 24: 113-122. http://dx.doi.org/10.5840/philtopics199624219

Mele, Alfred R. 1995. *Autonomous Agents*. New York: Oxford University Press.

Pereboom, Derek. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.

Pink, Thomas. 2019a. *Self-determination: The Ethics of Action*. Oxford: Oxford University Press, first published 2016.

Pink, Thomas. 2019b. "Freedom, Power and Causation." *Organon F*, 26(1): 141–168. https://doi.org/10.31577/orgf.2019.26109

Pinto, Anabela. 2022. *An Evolutionary Approach to the Adaptive Value of Belief* in *Evolutionary Psychology Meets Social Neuroscience*, edited by Rosalba Morese et al, Ch.2. London: IntechOpen.

Putnam, Hilary. 1985. *Mathematics, Matter and Method*. Cambridge: Cambridge University Press, first published 1975.

Shanahan, Murray. 2010. *Embodiment and the Inner Life: cognition and consciousness in the space of possible minds*. Oxford: Oxford University Press.

Sober, Elliott. 2008. *Evidence and Evolution: The Logic behind the Science*. Cambridge: Cambridge University Press.

Søvik, Atle O. 2018. *Free will, Causality and the Self*. Berlin: De Gruyter.

Anil, Seth. 2021. *Being You*. London: Faber & Faber.

Steward, Helen. 2014. *A Metaphysics for Freedom*. Oxford: Oxford University Press, first published 2012.

Strawson, Galan. 1986. *Freedom and Belief*. Oxford: Clarendon Press.

Tononi, Giulio et al. 2004. "An Information Integration Theory of Consciousness." *BMC Neuroscience*, 5:42. https://doi.org/10.1186/1471-2202-5-42

Tononi, Giulio et al. 2023. "Only What Exists Can Cause: An Intrinsic Powers View of Free Will." arXiv. https://doi.org/10.48550/arXiv.2206.02069

Wisniewski, David et al. 2019. "Free Will Beliefs Are Better Predicted by Dualism than Determinism Beliefs across Different Cultures." *PLoS ONE*, 14(9). https://doi.org/10.1371/journal.pone.0221617

Wittgenstein, Ludwig. 1976. *Philosophical Investigations*. Oxford: Basil Blackwell, first published 1953.