PSYCHOLOGY MISCELLANY

No.205 – August 2024

Kevin Brewer

ISSN: 1754-2200

orsettpsychologicalservices@phonecoop.coop

This document is produced under two principles:

1. All work is sourced to the original authors. The images are all available in the public domain (most from http://commons.wikimedia.org/wiki/Main_Page). You are free to use this document, but, please, quote the source (Kevin Brewer 2024) and do not claim it as you own work.

This work is licensed under the Creative Commons Attribution (by) 3.0 License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/ or send a letter to Creative Commons, 171 2nd Street, Suite 300, San Francisco, California, 94105, USA.

2. Details of the author are included so that the level of expertise of the writer can be assessed. This compares to documents which are not named and it is not possible to tell if the writer has any knowledge about their subject.

This document is presented for human readers.

Kevin Brewer BSocSc, MSc

An independent academic psychologist, based in England, who has written extensively on different areas of psychology with an emphasis on the critical stance towards traditional ideas.

A complete listing of his writings at http://psychologywritings.synthasite.com/. See also material at https://archive.org/details/orsett-psych.

CONTENTS

1. UNDERSTANDING CONSCIOUSNESS WITH POPULAR AND ACADEMIC LITERATURE

1.1. INTRODUCTION

In a special feature in the "New Scientist", Young et al (2021) outlined ten key questions about consciousness (which they tried to answer briefly):

1. What is consciousness?
2. How many states of consciousness do humans have?
3. Can physics explain consciousness?
4. What is consciousness like in other animals?
5. When did consciousness evolve (appendix 1A)?
6. How would we know if a machine was conscious?
7. What is consciousness for?
8. Is consciousness detectable in the brain?
9. Is the universe conscious?
10. Does consciousness make reality?

1.2. OVERVIEW

Neuroscience is helping in the understanding of the working of the physical brain. But how does this relate to who we are and our experience of being?

The idea of networks and connectivity in the brain has helped in the understanding of cognitive functions, particularly as neuroimaging has allowed the study of living brains. Bertolero and Bassett (2022) used the analogue of an orchestra playing a complex piece of

music.

The brain has also been viewed as separated into modules. Each module has a certain function "just as every family of instruments plays a role in the symphony" (Bertolero and Bassett 2022 p9). On top of the general patterns (of modules interconnected hubs) are "slight individual variations in the way our neural circuits are wired" (Bertolero and Bassett 2022 p9).

From a neuroscience viewpoint, the function and wiring of the brain describe the "who you are". "Put simply, your thoughts, feelings, quirks, flaws and mental strengths are all encoded by the specific organisation of the brain as a unified, integrated network. In sum, it is the music your brain plays that makes you you" (Bertolero and Bassett 2022 p12).

Seth (2021) talked of the "real problem" of consciousness: "how to explain, predict and control the various properties of consciousness in terms of physical processes in the brain and body" (p45). He distinguished between "conscious level (how conscious you are, as in the difference between general anaesthesia and normal wakeful awareness), conscious content (what you are conscious of) and conscious self (the experience of 'being you' - or, 'being me')" (Seth 2021 pp46-47) [1].

Levine (1983) used the phrase "explanatory gap" that needed filling in terms of the relationship between the physical brain and conscious experience.

A reductionist approach to disease, say, seeks the "lower" level of analysis (eg: gene). "Causal emergence" is an alternative which asserts that "the real cause of the disease emerges only at a higher level" (Ball 2022 p44) or that "a causal relationship is stronger at the macro-scale" (Comolatti and Hoel 2022 p1). Applying this to consciousness, a reductionist approach studies the individual neurons, say, while causal emergence would argue that consciousness emerges from the complex networks of neurons.

The sense of interoception, which monitors internal body signals like heart rate and pain, is now felt to play a major role in consciousness. "The brain on its own isn't enough to generate subjective experience... Without

[1] The experience of self as subject ("I") can be distinguished from as object ("Me") (Mead 1934). The former is "when we are minimally aware of ourselves and focus instead on the surrounding world" (Fernyhough and Borghi 2023 p1188), while the self as object is when the self is the focus of our thoughts. The self-as-subject experiences were reported more frequently in a study with university students, for instance (Morin 2022).

the body, the self simply wouldn't exist. 'Just as the notion of ''car'' exists only if a certain number of components are present and interacting with each other' [neuroscientist Catherine Tallon-Baudry]" (Williams 2023 p41).

Some take the idea further and suggest that "internal body signals aren't just involved in consciousness - they are consciousness. 'People continue talking about consciousness as the great mystery that will be revealed by understanding the brain, and that's wrong... It's not about the brain, it's about what the brain achieves with the interoceptive system in the body' [Antonio Damasio]" (Williams 2023 p42). This "body-centred" view of consciousness sees the brain as having an operational role in consciousness, but it "still doesn't explain how physical processes turn into that feeling of being 'you'" (Williams 2023).

Current theories of consciousness can be divided into four broad theoretical approaches (Seth 2021):

i) "Higher order" theories - "a mental state is conscious when another mental state - higher up in a hierarchy - says that it is" (Seth 2021 p46).

ii) "Global workspace" theories - "mental states are conscious when they are broadcast widely throughout the brain, so that they can be used to flexibly guide behaviour" (Seth 2021 p46).

iii) "Integrated information" theory - Consciousness depends on "the ability of a system to generate integrated information" (Seth 2021 p46).

iv) "Predictive processing" - Perception "isn't a passive registration of an external reality. It is an active construction, a kind of 'controlled hallucination', in which the brain's best guesses are tied to the world - and the body - through a continuous process of prediction error minimisation" (Seth 2021 p47).

1.3. QUALIA

Kukkonen (2023) asked: "Do I experience something as warm simply because my perceptual apparatus responds to a source of heat? What about the particular subjective dimension of our experience of different kinds of

warmth?" (p2). Clark (2018) used the term "strange inversion" to describe such subjective experience (or "qualia"). These "'occur when things work in ways that turn received wisdom upside down' [Clark 2018]. In the case of qualia..., the strange inversion consists in the fact that '[w]hat appears to be a property of the world itself here turns out [...] to be a property of the observer' [Clark 2018]. In other words, an object would not make us feel and act in certain ways because it is warm (or red, or sweet etc), but it would be inferred to be warm (or red, or sweet etc) because it makes us feel and act in certain ways" (Kukkonen 2023 p2).

Clark (2019) linked subjective experience to perception of the world: Just as we perceive objects, like dogs, cats, and tables, we perceive "the presence of rather puzzling 'qualitative features' such as cuteness, brownness, or annoyingness of the dog" (quoted in Kukkonen 2023).

1.4. MEASUREMENT

The idea of a "consciousness meter", though sounding like science-fiction, could help in cases of coma (and related conditions) to know if the individual is experiencing anything. Technological developments have offered the possibility of such a device (Koch 2022a).

The electroencephalogram (EEG) was the first attempt to measure brain waves and thus consciousness. Developed by Hans Berger in the 1920s, the EEG measures minute voltage fluctuations in electrical activity across the surface of the brain (10 to 100 micro-volts). Modern versions can use over 200 electrodes on the surface of the skull. A signature of consciousness can be observed ("low voltage, rapid up-and-down fluctuating waves that are desynchronised rather than in lockstep across the skull"; Kock 2022 p17).

EEG-based data have been used in the "Integrated Information Theory" (IIT) of consciousness (eg: Tononi et al 2016). EEG responses across the brain in different areas and in different individuals can be converted into the "perturbational complexity index" (PCI) (Koch 2022a) [2]. Casarotto et al (eg: 2016) measured EEG across the brain of 102 non-brain-injured participants and 48 still responsive and awake brain-injured patients receiving transcranial magnetic stimulation. Based on comparison of

[2] IIT can be simplified to "something has a higher level of consciousness if the interactions between its components yield more information than when reduced to just its components. In other words, the whole is greater than the sum of its parts" (Wilson 2023 p8).

different states, a PCI score of 0.31 was calculated as the threshold of consciousness - if above that number, the individual was conscious, while below it the individual was unconscious (Koch 2022a). However, subsequent research has found that the distinction is not so clear-cut with individuals with "unresponsive wakefulness syndrome" (UWS), for instance (where sufferers show sleep-wake cycles and have basic processes like normal heart rate, but no intentions) (Koch 2022a).

Other recent research has used functional magnetic resonance (fMRI) scans. For example, Nemirovsky et al (2023) scanned seventeen people when awake, mildly sedated, unconscious/deep anaesthesia, and in recovery from the anaesthesia, and found differences in certain brain networks - the fronto-parietal and dorsal attention networks. The PCI score was calculated for these areas, and it was lower in deep anaesthesia than awake (Wilson 2023).

Russell Hurlburt in the 1970s introduced the idea of a beeper study for introspection. Randomly a beeper would sound during the day and study participants have to record their current inner experience (eg: Heavey and Hurlburt 2008; 30 US students were beeped randomly six times per day on three days; table 1.1). This has been called "descriptive experience sampling" (DES) (Douglas 2023).

PHENOMENA	EXAMPLE	OVERALL FREQUENCY IN REPORTS (%) *
Inner speech	"Susan" saying to herself mentally: "I've got to get to class".	26
Inner seeing	"Paul" imaginally seeing the face of best friend.	34
Unsymbolised thinking	"Adam" watching two men carrying bricks and wondering whether they would drop any.	22
Feeling	"Courtney" was angry.	26
Sensory awareness	"Harriet" feeling the cold breeze on the left cheek on a windy day.	22

(* "Most participants had one form of inner experience predominate; 22 of the 30 participants had at least one of the five common phenomena occurring in 50% or more of their samples"; Heavey and Hurlburt 2008 p803)

(Based on tables 1 and 2 Heavey and Hurlburt 2008)

Table 1.1 - Inner experience reported in Heavey and Hurlburt (2008) study.

What people report varies, not just in content, but in what is perceived as thought, such that five common phenomena have been distinguished - inner speech (appendix 1B), inner seeing, feelings or emotions, sensory awareness (eg: the stone in my shoe), and unsymbolised thinking (explicit thinking without language) (Douglas 2023).

But these types of thoughts overlap (Douglas 2023).

Self-report questionnaires is another method of study, like the "Internal Representations Questionnaire" (IRQ) (Roebuck and Lupyan 2020). This was designed to measure "internal verbalisation" (ie: "the tendency to experience thoughts in the form of language"; pp2053-2054), or more specifically, "to measure people's propensity to use internalised language in different situations that do not involve communication with other people. These include using language as a retrieval cue for autobiographical memories, for cueing a visualisation of a scene (eg: visualising a beach by internally using the word 'beach'), and talking with oneself when trying to work out a problem" (Roebuck and Lupyan 2020 p2054).

The construction of the IRQ began by listing all relevant statements based on the previous literature (n = 81), which were presented to 180 US students. Based on their responses, 60 items were chosen, and administered to a sample of 222 adults recruited from "Amazon's Mechanical Turk". Factor analysis of the responses using a five-point scale ("strongly disagree" to "strongly agree") produced four underlying factors, and 36 statements for the final version of the IRQ (table 1.2)

- Visual Imagery (10 items) - eg: "I often enjoy the use of mental pictures to reminisce"; "I can close my eyes and easily picture a scene that I have experienced".

- Internal Verbalisation (12 items) - eg: "My inner speech helps my imagination"; "I think about problems in my mind in the form of a conversation with myself".

- Orthographic Imagery (6 items) - eg: "I see words in my 'mind's eye' when I think"; "When I hear someone talking, I see words written down in my mind".

- Representational Manipulation (8 items) - eg: "In school, I had no problems with geometry"; "I can easily imagine the sound of a trumpet getting louder".

(Source: table 2 p2056 Roebuck and Lupyan 2020)

Table 1.2 - Example of statements of the IRQ.

1.5. INTRA-CRANIAL STIMULATION

"Intra-cranial brain stimulation highlights the daily miracle of the brain's water changing into the wine of consciousness" (Kock 2022b p31). It involves the direct electrical stimulation of areas of the surface of the brain [3], which produce spontaneous thoughts, feelings and memories in conscious patients.

Fox et al (2020) reported the intra-cranial electrical stimulation (iES) of 1537 sites of the cortex of 67 epileptic patients undergoing open-brain surgery (between 2008 and 2018 at the Stanford University Medical Centre) [4] [5]. Comments included: "You look down at yourself lying in bed from above but see only your legs and lower trunk", and "You see little white dots on a black background, as if looking up at the stars at night" (quoted in Koch 2022b).

1.6. CONSCIOUSNESS AND PHYSICS

"Modern neuroscience has left little room inside the brain for an immaterial soul. Instead, physicalism reigns - the idea that everything in nature must be derived from the basic stuff of physics. It follows that consciousness must somehow emerge out of particles, strings, information or whatever you take as fundamental" (Lawton 2022 p38).

This has been called the "brute identity theory" (ie: conscious states are simply states of the brain). Carroll (2021), however, made a distinction: "It's not that conscious states 'are' states of the brain; it's that certain states of the brain correspond to certain conscious states".

In relation to the "hard problem of consciousness" (appendix 1C) (Chalmers 2007) - how the physical processes of the brain give rise to the subjective experience of consciousness - physicalism sees this as a "red herring", as when we know enough about the physical brain consciousness will be explained (Lawton 2022).

An alternative is "panpsychism", which suggests that "some form of consciousness, however fragmentary, is an intrinsic property of matter" (Lawton 2022 p40). But this view is challenged. For example, philosopher Eleanor Knox argued: "We should be sceptical of attempts to amalgamate

[3] This technique was pioneered by Wilder Penfield (eg: Penfield and Rasmussen 1950).

[4] The patients underwent iES as part of routine to map seizure areas (Fox et al 2020).

[5] "To control for demand characteristics and false positive reports, a total of 116 sham stimulations were also delivered" (Fox et al 2020 p1040).

lots of little bits of micro-consciousness to create complex consciousness" (quoted in Lawton 2022).

Another view is "illusionism", which basically sees consciousness as an illusion (ie: it does not exist) (Seth 2021). While "passive mentalism" is the idea that there are purely mental aspects that have no physical dynamics (Carroll 2021).

Carroll (2021) observed: "We have a much better understanding of physics than we do of consciousness" (p16). Physics provides evidence for physical events having purely physical causes. However, Carroll (2021) continued, "our current knowledge of physics should make us sceptical of hypothetical modifications of the known rules, and that without such modifications it's hard to imagine how intrinsically mental aspects could play a useful explanatory role" (p16).

Carroll (2021) ended: "The most promising route to understanding consciousness is likely to involve further neuroscientific insights and a more refined philosophical understanding of weak emergence, rather than rethinking the fundamental nature of reality".

Smolin et al (2021) argued that the problems of qualia and conscious awareness are linked to other deep problems in quantum theory, and the nature of time.

1.7. MACRO OR MICRO

Hoel et al (2013) began: "In science, it is usually assumed that, the better one can characterise the detailed causal mechanisms of a complex system, the more one can understand how the system works. At times, it may be convenient to resort to a 'macro'-level description, either because not all of the 'micro'-level data are available, or because a rough model may suffice for one's purposes. However, a complete understanding of how a system functions, and the ability to predict its behaviour precisely, would seem to require the full knowledge of causal interactions at the micro level. For example, the brain can be characterised at a macro scale of brain regions and pathways, a meso scale of local populations of neurons such as mini-columns and their connectivity, and a micro scale of neurons and their synapses. With the goal of a complete mechanistic understanding of the brain, ambitious programs have been launched with the aim of modelling its micro scale" (p19790).

The assumption is that "when the properties of

micro-level physical mechanisms of a system are fixed, so are the properties of all its macro levels — a relation called 'supervenience' [Davidson 1980]. In turn, this relation is usually taken to imply that the micro mechanisms do all of the causal work, ie: the micro level is causally complete. This leaves no room for any causal contribution at the macro level; otherwise, there would be 'multiple causation' [Kim 1993]" (Hoel et al 2013 p19790).

The alternative to this reductionism of the micro-level focus is emergence or the idea that the "whole is greater than the sum of the parts". Hoel et al (2013) provided support for this latter position using a simple simulated system. They argued for "the intuitive idea that, to find out how a system works, one should find the 'differences that make [most of] a difference' to the system itself [Tononi 2012]. It also suggests that complex, multi-level systems such as brains are likely to work' at a macro level because, in biological systems, selectional processes must deal with unpredictability and lead to degeneracy [Tononi et al 1999]" (Hoel et al 2013 p19795).

1.8. PHILOSOPHICAL THOUGHTS

Horgan (2022) asked the philosophical question, "How do I know I'm not the only conscious being in the universe?". He called this the "solipsism problem". Though this seems like a frivolous question, "each of us is sealed in an impermeable prison cell of subjective awareness... You experience your own mind every waking second, but you can only infer the existence of other minds through indirect means" (Horgan 2022 p34).

1.9. MISCELLANEOUS ISSUES

"Brain organoids" are tiny structures grown from human stem cells, and they are common in petri-dish-based laboratory studies. Trujillo et al (2019) reported electrical activity from such an organoid that resembled that seen in premature babies, and this type of brain-wide co-ordinated electrical activity is often used as a definition of consciousness (Reardon 2022).

This raises the question as to whether laboratory-grown brains are or could become conscious. But one problem is the lack of agreed operational definition of consciousness, and it depends upon the theory of

consciousness held. "Creating a conscious system might be a whole lot easier than defining it" (Reardon 2022 p22).

Developmental biologist Madeline Lancaster commented: "If you thought a fly was conscious, it's conceivable that an organoid could be" (quoted in Reardon 2022). While Alysson Muotri (of Trujillo et al 2019) stated this pragmatic view: "We work with animal models that are conscious, and there are no problems... We need to move forward, and if it turns out they become conscious, to be honest, I don't see it as a big deal" (quoted in Reardon 2022).

1.10. FREE WILL AND NEUROSCIENCE

The growth of neuroscience and its technology has led "toward the thesis that free will is pure illusion and that the principle of causation in all its rigour leads inexorably to the rejection of a concept of free will likely to contribute to an understanding of human behaviour" (Feltz et al 2020 p1).

A key set of experiments for this position comes from Libet (1985) (table 1.3), which recorded electrical activity in the brain before an individual consciously decided to move their hand, say. Specifically, a readiness potential recorded half a second before, and conscious intention a quarter of a second before the action (Nahmias 2022) (appendix 1D) [6] [7]. The cause of the behaviour is outside of consciousness, and the individual constructs their decision as free will after the event, according to Wegner (2002) [8]. "In this context, free will proper has no place and is qualified as a 'necessary illusion' in that, while we must act as if our decisions were efficacious, this efficacy is fictive on account of the principle of total determinism" (Feltz et al 2020 p2). Mele (2009) argued that the Libet (1985) experiment "falls well short of justifying the claim that free will is an illusion" (Mele 2020 p83).

Nahmias (2022) used the term "willusionists" to describe those who see free will as an illusion, while Mele (2018) favoured a "decision-focused sceptical argument" (DSA).

[6] While, using electrodes planted in the brain, Fried et al (2011) recorded neuronal activity in the supplementary motor area 700 ms before the conscious decision to press a button.

[7] Trevena and Miller (2010) used similar methods, but drew different conclusions.

[8] This has been described as epiphenomenalism - "This is the theoretical thesis that seemingly causally relevant conscious processes, such as intention formation or decisions, do not play a causal role in the initiation of the corresponding action" (Bonicalzi 2020 p125).

- Participants watched a clock and remembered the time when they consciously decided to move (called "W"), while the actual muscle movement ("M") was measured by an electromyogram, and brain activity an electroencephalograph (EEG) system ("readiness potential"; RP).

- Non pre-planned actions: RP 550 ms before M and 350 ms before W.
- Planned actions: RP 1000 ms before M (Bonicalzi 2020).

- Stating the time of W is really "the timing at which they consciously experience they wanted to make a movement. This moment of awareness seemingly corresponds to a second-order state or meta-state (ie: the consciousness of the wish to make a movement), rather than to a first-order intentional state (ie: the wish to make a movement)" asserted Bonicalzi (2020 p127).

- Khalighinejad et al (2018), for example, overcame the problem of telling participants to perform an intentional action by giving them a task of waiting for dots to move on a computer screen which they could skip if they wanted by pressing a button. "Voluntary actions were then operationalised as self-initiated skip responses while waiting for the display of dots to move coherently towards the left or the right. This way, the experimental paradigm was able to elicit the performance of intentional, or at least voluntary, actions without artificially probing participants to act" (Bonicalzi 2020 p128).

- Haggard and Eimer (1999) found that the onset of RP could vary independently of W. "In particular, trials where participants show an early W were characterised by a late onset of the RP, compared to trials characterised by a late W. A better candidate for *causing* W might be the lateralised readiness potential (lrp), an increase in the electrical negativity in the area contra-lateral to the subsequent bodily movement, and reflecting the preparation of a specific movement after the action selection is made" (Bonicalzi 2020 p128).

Table 1.3 - Libet (1985).

Mele (2020) made the distinction between distal and proximal intentions [9]. The latter was studied by Libet (1985), and is the intention to act now. Distal intention is a longer term decision - eg: to go to the park at the weekend.

Mele (2020) outlined two key challenges to Libet (1985) and other similar studies:

a) When do participants make their decisions? This

[9] Bratman (1987) distinguished between future-directed intention and present-directed intention (Bonicalzi 2020).

is a challenge to the method where participants have say "now", for example, when they make a conscious decision to make their hand, as well as to the theoretical distinction between conscious and unconscious causes of behaviour.

Mele (2020) explained: "Why did I say 'now!' exactly when I did? On any given trial, I had before me a string of equally good moments for a 'now!' – saying, and I arbitrarily picked one of the moments. But what led me to pick the moment I picked? The answer offered by Schurger et al [2012] is that random noise crossed a decision threshold then. And they locate the time of the crossing very close to the onset of muscle activity – about 100 ms before it... They write: 'The reason we do not experience the urge to move as having happened earlier than about 200 ms before movement onset [referring to Libet's partipants' reported... time] is simply because, at that time, the neural decision to move (crossingvthe decision threshold) has not yet been made'" (p90).

b) When do participants reach the point of no return for an action? Libet's (1985) answer would be before the conscious decision. But Schultze-Kraft et al (2016), using brain imaging, found evidence of the ability to cancel an "unconscious" command (Mele 2020).

Balaguer (2009) argued that "individuals are truly acting freely only when they make choices between options about which they are authentically torn, such as life-changing decisions about different job opportunities... In contrast, the decision-making context associated to neuroscientific paradigms looks more similar to a situation where I am in front of a shelf in a supermarket and have to repeatedly choose between identical boxes of cereals - something I can do almost automatically or at least without the vigilant monitoring of conscious mental states. A related interpretative problem might arise if the type of actions experimental subjects make are so low-level that they become automatic or absent-minded: As a consequence, the data might be scarcely informative regarding the neural bases of conscious intentions" (Bonicalzi 2020 p126).

Taking a different approach, Feltz et al's (2020) saw "intentional action in human beings is a function of both the initial conditions of the distributed neural networks that are involved (that is, the 'brain state': emotions, physiological state, autobiographical and implicit memory) and the circumstances and events at some

moment (rest, activity, social interactions, etc). The expression 'free will' is used in order to describe this situation of interaction between an agent whose nervous system reacts to a prior set of events according to both its state at that point and according to an intentional logic. Intentional action refers explicitly to an operation that involves the capacity to represent a future state of the world" (Feltz et al 2020 p2). Language is key in this operation.

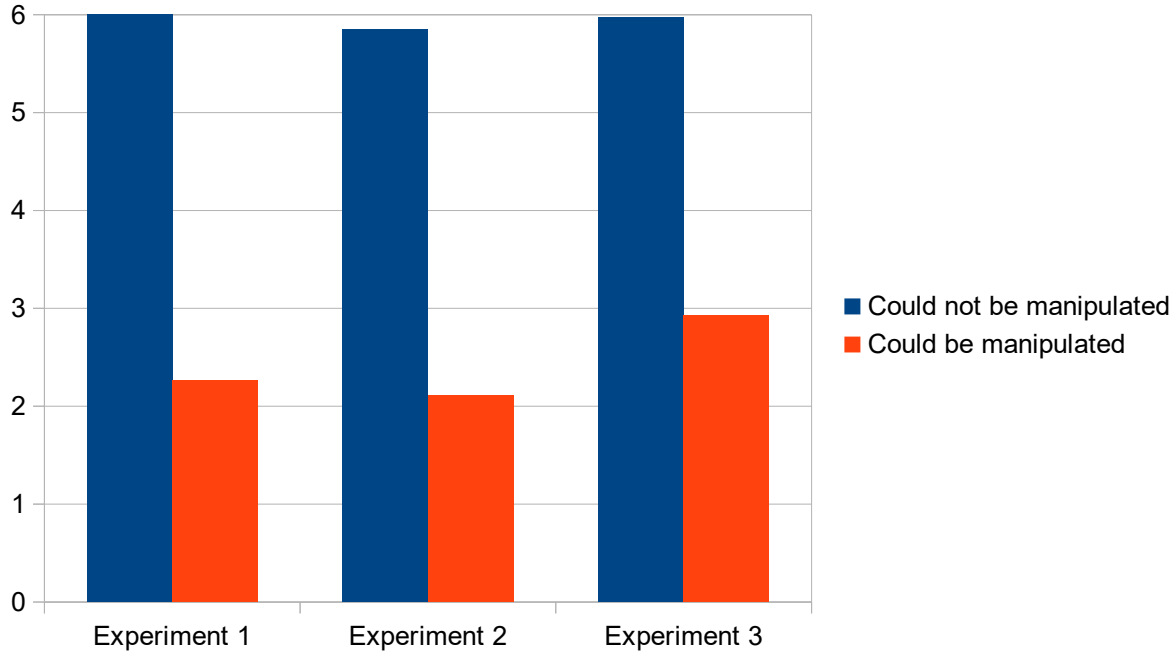
Frankfurt (1977) produced a hierarchical explanation of of conscious mental states in relation to the concept of a person. "Intrinsic to the concept of a person is the capacity to identify (with) higher order mental states (ie: the desire to have the desire to drink water) ideally exercising control over the lower level mental states (ie: the desire to drink water) that are ultimately responsible for the agent's behaviour. In the sequence leading to action production, the authority of conscious mental states is thus directly rooted in the agent's psychological make-up: For an agent to qualify as a person, she must be able to identify (with), or endorse, the mental states from which her actions stem" (Bonicalzi 2020 p123). Bonicalzi (2020) used the term "intentional action control" to describe the "over-riding force of mental states over bodily actions" (p124).

Nahmias et al (2014) undertook three experiments on everyday understanding of free will. In the first experiment, 278 US undergraduates were presented with this scenario: "Neuroscientists can use brain scanners to detect all the activity in a person's brain and use that information to predict with 100% accuracy every single decision a person will make before the person is consciously aware of their decision" (p504). A character called "Jill" experiences this situation in a number of different ways (eg: voting in an election), and there were two versions of the scenario, either the scientists could or could not manipulate behaviour. Participants perceived "Jill" as having free will when the scientists could not manipulate her behaviour. The second experiment was a replication with 213 more students.

The third experiment used a similar idea, but "a colony of mindreaders" (not neuroscientists) knew future behaviour and could/could not manipulate it. Over 90% of 196 US students agreed that "Jill" had free will when the mindreaders could predict but not manipulate her actions (figure 1.1).

"Most people seem to understand free will in a way that is not threatened by perfect prediction based on

neural information, suggesting that they believe that just because 'my brain made me do it', that does not mean that I didn't do it of my own free will" (Nahmias et al 2014 p502).



(Data from Nahmias et al 2014)

Figure 1.1 - Mean rating of Jill having free will in three experiments where she could be manipulated or not.

1.11. APPENDIX 1A - EVOLUTION OF CONSCIOUSNESS

In terms of the evolution of consciousness, the question is whether there is a single common ancestor or that consciousness evolved independently in different lines of the animal kingdom (Young et al 2021).

The single common ancestor idea (eg: Jablonka and Ginsburg 2019) "hinges on a type of thinking known as unlimited associative learning, which involves learning about novel stimuli and linking them together even if time has passed between experiencing two stimuli" (Young et al 2021 p39). Simple learning developing into complex learning requires consciousness to direct attention in the environment. Unlimited associative learning is seen in many ancient species, so consciousness may have appeared in early vertebrates around 530 million years ago (Young et al 2021).

The independent evolution of consciousness in

different phyla is based on the argument that the same function is performed by different brain areas in different species (eg: conscious visual perception) (Young et al 2021).

1.12. APPENDIX 1B - INNER SPEECH

"Inner speech" (IS) is a "significant feature" of subjective experience, variously described as inner (or internal) monologue, inner (or internal) dialogue, inner voice, covert (or silent) self-talk, and internal narrative, for instance (Fernyhough and Borghi 2023).

IS has been found to help in focused attention tasks, and to enhance cognitive flexibility, and memory (Fernyhough and Borghi 2023).

The "Varieties of Inner Speech Questionnaire" (VISQ) (McCarthy-Jones and Fernyhough 2011) was developed to measure five dimensions of IS - dialogicality (IS as a dialogue), condensation (use of full sentences or not), the presence of other voices, evaluation/critical content, and positive/regulatory (eg: motivated self-talk) (Fernyhough and Borghi 2023).

1.13. APPENDIX 1C - HARD AND EASY PROBLEMS

Chalmers (2007) asserted: "There is not just one problem of consciousness. 'Consciousness' is an ambiguous term, referring to many different phenomena. Each of these phenomena needs to be explained, but some are easier to explain than others" (p225). This has led to the distinction between "hard" and "easy" problems. "The easy problems of consciousness are those that seem directly susceptible to the standard methods of cognitive science, whereby a phenomenon is explained in terms of computational or neural mechanisms. The hard problems are those that seem to resist those methods" (Chalmers 2007 p225).

Chalmers (2007) outlined seven "easy" problems, including the focus of attention, the difference between wakefulness and sleep, and the deliberate control of behaviour. The "hard" problem of consciousness is "the problem of experience": "When we think and perceive, there is a whir of information-processing, but there is also a subjective aspect. As Nagel (1974) has put it, there is 'something it's like' to be a conscious organism. This subjective aspect is experience" (Chalmers 2007 p226). Linked to this is the problem of how physical

processing gives rise to the subjective experience.

Chalmers (2007) recommended that the "hard" problem be called "consciousness", and the "easy" problems "awareness".

1.14. APPENDIX 1D - SOON ET AL (2013)

Similar findings to Libet (1985) comes from Soon et al (2013) using neuroimaging.

Participants - Seventeen adults.

Procedure - Numbers were presented on a computer screen. "Participants spontaneously and freely decided to perform either of two abstract intentions: adding or subtracting" (Soon et al 2013 p6220). As soon as participants "first became conscious of the spontaneous urge to perform either addition or subtraction" (p6220), they made it known. This study took place during functional magnetic resonance image (fMRI) scanning.

Findings - Neural activity in the medial prefrontal and parietal cortex up to four seconds before the conscious decision was reported.

Comment - Previous research had focused on unconscious activity prior to conscious motor activity, whereas this study showed brain activity prior to "a non-motor abstract mental operation" (Soon et al 2013 p6219).

John-Dylan Haynes (of Soon et al 2013) stated that "it seems that the brain is making the decision before the person" (quoted in Nahmias 2022).

1.15. REFERENCES

Balaguer, M (2009) Free Will as a Scientific Problem Cambridge, MA: MIT Press

Ball, P (2022) What causes what? New Scientist 28th May, 44-47

Bertolero, M & Bassett, S (2022) How matter becomes mind Scientific American 31, 1 (Special edition: Secrets of the Mind), 6-13

Bonicalzi, S (2020) Actions and intentions. In Feltz, B et al (eds) Free Will, Causality and Neuroscience Leiden: Brill

Bratman, M.E (1987) Intentions, Plans and Practical Reason Cambridge: Cambridge University Press

Carroll, S.M (2021) Consciousness and the laws of physics Journal of Consciousness Studies 28, 9-10, 16-31

Casarotto, S et al (2016) Exploring the neurophysiological correlates of loss and recovery of consciousness: Perturbational complexity. In Monit, M & Sannita, W (eds) Brain Function and Responsiveness in Disorders of Consciousness New York: Springer

Chalmers, D (2007) The hard problem of consciousness. In Velmans, M & Schneider, S (eds) The Blackwell Companion to Consciousness Oxford: Blackwell

Clark, A (2018) Strange inversions: Prediction and the explanation of conscious experience. In Huebner, B (ed) The Philosophy of Daniel Dennett Oxford: Oxford University Press

Clark, A (2019) Consciousness as generative entanglement Journal of Philosophy 116, 12, 645-662

Comolatti, R & Hoel, E (2022) Causal emergence is widespread across measures of causation arXiv (https://arxiv.org/abs/2202.01854)

Davidson, D (1980) Mental events. In Block, N (ed) Readings in Philosophy of Psychology Cambridge, MA: Harvard University Press

Douglas, K (2023) How are you thinking? New Scientist 22nd July, 32-35

Feltz, B et al (2020) Introduction. In Feltz, B et al (eds) Free Will, Causality and Neuroscience Leiden: Brill

Fernyhough, C & Borghi, A (2023) Inner speech as language process and cognitive tool Trends in Cognitive Science 27, 12, 1180-1193

Fox, K.C.R et al (2020) Intrinsic network architecture predicts the effects elicited by intra-cranial electrical stimulation of the human brain Nature Human Behaviour 4, 10, 1039-1052

Frankfurt, H (1977) The problem of action Philosophical Quarterly 15, 2, 157-162

Fried, I et al (2011) Internally generated pre-activation of single neurons in human medial frontal cortex predicts volition Neuron 69, 548-562

Haggard, P & Eimer, M (1999) On the relation between brain potentials and awareness of voluntary movements Experimental Brain Research 126, 1, 128-133

Heavey, C.L & Hurlburt, R.T (2008) The phenomena of inner experience Consciousness and Cognition 17, 3, 798-810

Hoel, E.P et al (2013) Quantifying causal emergence shows that macro can beat micro Proceedings of the National Academy of Sciences, USA 110, 49, 19790-19795

Horgan, J (2022) How do I know I'm not the only conscious being in the universe? Scientific American 31, 1 (Special edition: Secrets

of the Mind), 32-35

Jablonka, E & Ginsburg, S (2019) The Evolution of the Sensitive Soul: Learning and the Origins of Consciousness Cambridge, MA: MIT Press

Khalighinejad, N et al (2018) Precursor processes of human self-initiated action Neuroimage 165, 35-47

Kim, J (1993) Supervenience and Mind: Selected Philosophical Essays Cambridge: Cambridge University Press

Koch, C (2022a) How to make a consciousness meter Scientific American 31, 1 (Special edition: Secrets of the Mind), 14-19

Koch, C (2022b) The brain electric Scientific American 31, 1 (Special edition: Secrets of the Mind), 26-31

Kukkonen, K (2023) Designs on consciousness: Literature and predictive processing Philosophical Transactions of the Royal Society B 379, 20220423

Lawton, T (2022) Consciousness in the cosmos New Scientist 2nd April, 38-42

Levine, J (1983) Materialism and qualia: The explanatory gap Pacific Philosophical Quarterly 64, 354-361

Libet, B (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action Behavioural and Brain Sciences 8, 529-566

McCarthy-Jones, S & Fernyhough, C (2011) The variation of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults Consciousness and Cognition 20, 4, 1586-1593

Mead, G.H (1934) Mind, Self, and Society from the Standpoint of a Social Behaviourist Chicago: University of Chicago Press

Mele, A (2009) Effective Intentions: The Power of Conscious Will New York: Oxford University Press

Mele, A (2018) Free will and consciousness. In Jacquette, D (ed) Bloomsbury Companion to the Philosophy of Consciousness London: Bloomsbury

Mele, A (2020) Free will and neuroscience: Decision times and the point of no return. In Feltz, B et al (eds) Free Will, Causality and Neuroscience Leiden: Brill

Morin, A (2022) Self-reported inner speech illuminates the frequency and content of self-as-subject and self-as-object experiences Psychology of Consciousness: Theory, Research, and Practice 9, 1, 93-100

Nagel, T (1974) What is it like to be a bat? Philosophical Review 4, 435-450

Nahmias, E et al (2014) It's OK if "my brain made me do it": People's intuition about free will and neuroscientific prediction Cognition 133, 502-516

Nahmias, E (2023) Why we have free will Scientific American 31, 1 (Special edition: Secrets of the Mind), 44-47

Nemirovsky, I.E et al (2023) The implementation of integrated information theory in resting-state fMRI Communications Biology 6, article 692

Penfield, W & Rasmussen, T (1950) The Cerebral Cortex of Man: A Clinical Study of Localisation of Function London: Macmillan

Reardon, S (2022) Can lab-grown brains become conscious? Scientific American 31, 1 (Special edition: Secrets of the Mind), 20-25

Roebuck, H & Lupyan, G (2020) The Internal Representations Questionnaire: Measuring modes of thinking Behaviour Research Methods 52, 2053-2070

Schultze-Kraft, M et al (2016) The point of no return in vetoing self-initiated movement Proceedings of the National Academy of Sciences, USA 113, 1080-1085

Schurger, A et al (2012) An accumulator model for spontaneous neural activity prior to self-initiated movement Proceedings of the National Academy of Sciences, USA 109, 42 & E2904-2913

Seth, A (2021) Illuminating consciousness New Scientist 4th September, 45-48

Smolin, L et al (2021) Physics, time and qualia Journal of Consciousness Studies 28, 36-51

Soon, C.S et al (2013) Predicting free choices for abstract intentions Proceedings of the National Academy of Sciences, USA 110, 15, 6217-6222

Tononi, G (2012) Integrated information theory of consciousness: An updated account Archives Italiennes de Biologie 150, 2-3, 56-90

Tononi, G et al (1999) Measures of degeneracy and redundancy in biological networks Proceedings of the National Academy of Sciences, USA 96, 6, 3257-3262

Tononi, G et al (2016) Integrated information theory: From consciousness to its physical substrate Nature Reviews Neuroscience 17, 7, 450-461

Trevena, J & Miller, J (2010) Brain preparation before a voluntary action: Evidence against unconscious movement initiation Consciousness and Cognition 19, 447-456

Trujillo, C.A et al (2019) Complex oscillatory waves emerging from cortical organoids model early human brain network development Cell: Stem Cell 25, 4, 558-569

Wegner, D.M (2002) The Illusion of Conscious Will Cambridge, MA: MIT Press

Williams, C (2023) How to think about consciousness New Scientist 13th May, 41-42

Wilson, C (2023) A new way to know our minds New Scientist 22nd July, p8

Young, E et al (2021) Consciousness expanded New Scientist 10th July, 34-43

2. CRIMINALS VERSUS THE INSURANCE INDUSTRY

Becker (1968) described an economic model of criminal behaviour based on the net pay-off (ie: income minus expenditure), the probability of being caught and punished, and the severity of the punishment. Concentrating on the insurance industry, Baker and Shortland (2023) explained that "[I]nsurers can influence the cost of carrying out certain crimes, the payoffs to criminals, and the probability of being punished by interacting with insureds, governments, or third parties. Criminals can respond to these measures by reducing their activity, moderating the harm they inflict on insureds, or by innovating in turn to improve their expected utility. Criminal advances may trigger further innovations by insurers and vice versa: insurance and crime co-evolve" (pp183-184).

Insurers influence the insured to self-protect (eg: fitting alarm systems), third parties to create barriers to crime (eg: marking vehicles with unique identification numbers), and governments and law-makers in their decisions (eg: the use of CCTV systems in public places). But "the insurance industry does not have the incentive to suppress crime to the point where insurance becomes unnecessary" (Baker and Shortland 2023 p184) [10].

Baker and Shortland (2023) outlined five areas of crime and "insurers' efforts to keep various crimes insurable" (p185):

1. Vehicle theft

The first response to automobile theft was in 1913 in USA as insurers started a list of stolen vehicles which assisted the police. Criminals reacted by removing identification information, and insurers encouraged the development of technology to stop this or show it had happened. This is an example of the "arms race" between criminals and insurers in the last century.

2. Art theft

Since the 1950s increasing art theft has meant that insurers are motivated to support the security industry. However, "insider crime" defeats state-of-the-art security systems. Insurers responded with "rewards for

[10] Avraham and Porat (2022) talked of a "dark side of insurance".

information" in retrieving stolen art. "However, this resulted in an attractive criminal business model —or 'tickle'. Soon after a theft, someone would report having 'overheard' a conversation among complete strangers in an ill-lit bar revealing the location of the stolen paintings in — say — a station locker, dusty attic, or abandoned car and collect the reward. This minimal-risk route to divesting loot served both thieves and insurers, who much preferred paying rewards to reimbursing the full value... By setting the rewards (and making clear that only objects in good condition were eligible), insurers stayed in control of losses. If loss adjusters were (somehow) satisfied that the informant was not a criminal, no law was broken in the transaction. Yet, there was public and political disquiet about this cosy co-existence of insurance and crime" (Baker and Shortland 2023 pp187-188). Legislation in the UK in the late 1960s, for example, stopped such practices.

Another strategy of insurers was to make artworks "too hot to handle" (eg: via the creation of an art loss register). This was more effective for "high-end" artworks (Baker and Shortland 2023). Private art recovery also developed in the late 20th century along with "art recovery insurance". Baker and Shortland (2023) explained: "The market for insurance and art theft can, therefore, be considered in an unstable equilibrium. Iconic masterpieces are not attractive to steal as they are effectively unsaleable. For other high-end art, partial insurance limits losses, and some of the pay-outs are later salvaged through successful recoveries. The mid-market is either directly insurable or insurable for recovery. But high insurance premia and a slow recovery process mean that investment art is best kept in high-security storage areas — often in freeports — rather than displayed and enjoyed" (pp188-189).

3. Kidnap and hi-jack for ransom

"To be insurable, kidnapping and piracy must be predictable but rare, affordable to resolve, and essentially non-violent. This is far from the popular image of kidnapping, and indeed insurability is only achieved with significant (public and private) governance efforts. Kidnap- and hijack-for-ransom (K&R) insurance strikes a delicate balance between discouraging hostage-taking and facilitating the safe release of hostages... On the one hand, significant value must be attached to living hostages, so that criminals aim not to injure or

kill their victims during the abduction and keep them safe thereafter. Negotiations, payments, and releases must be managed as calmly and efficiently as possible to minimise the risk to life. On the other hand, offering ransoms, normalising extortion, and facilitating payments encourage kidnapping. Rising ransoms can create unstable dynamics: attracting (inexperienced) criminals into abductions and leading to (violent) kidnap hotspots" (Baker and Shortland 2023 p189).

K&R policies were first offered in the 1930s. Complex norms and processes have developed around such policies over the subsequent time, including the responsibility upon potential victims (and their employers), particularly in high-risk environments. "The objective is not to eliminate kidnapping, but to make it infrequent. Occasionally, an opportunist may strike lucky and sometimes there may be frictions in the (implicit) protection contracts. These incidents are handled by professional crisis responders, who are retained by insurers. Crisis response was created when K&R insurance was destabilised by massive ransom inflation in Latin America in the 1970s. Insurers realised that they needed to take control of ransom negotiations — or rather give advice to the insured on how to bargain with criminals to keep hostages safe" (Baker and Shortland 2023 p190).

Any equilibria are unstable. Companies with kidnapped employees, for example, may not follow insurers' protocols, and fast and "overgenerous" ransoms encourage "copycat crime". Insurers can respond by classifying an area a "war risk", which increases premia, but also incentivises companies and governments to reduce risks in such areas. Also "insureds buy the services of private security companies, who are engaged in their own technical arms race with kidnappers" (Baker and Shortland 2023 p191).

4. Ransomware

This is malware that encrypts computer data until a ransom is paid for a decryption key. Cyber-insurance for organisations to cover this situation appeared in the late 1990s.

Criminals in this field have two problems – making the encryption too difficult for security engineers to resolve, and the receipt of the ransom. The latter was aided by the development of cryptocurrencies in the early 21st century.

Another variable was that some criminals did not provide decryption keys after payment and/or the victims

still lost their data. "Insurers responded by helping their customers to streamline the resolution and recovery process by putting them directly in touch with professional breach responders. By pooling information, responders created incentives for hackers to provide aftercare in the resolution process. Customers were advised not to pay ransoms to ransomware groups with a track record of poor decryption success. Responders also collected and shared the decryption keys purchased by their customers to help other victims recover for free, thereby further raising the bar for criminals. Criminals thus had to work ever harder to create 'brands' of ransomware that were irreversible without a ransom, but reliably resolved once the payment was made" (Baker and Shortland 2023 p192).

Insurers have worked to reduce opportunities for unsophisticated criminals, developed engagement strategies for sophisticated criminals, but stayed away from state-sponsored cyber events (eg: exclusion clauses in cyber-insurance policies).

5. Credit card fraud

Insurance is not explicit here as in specific policies, but rather, banks and credit card companies implicitly insure customers by reimbursing them in fraudulent transactions.

The "arms race" is thus between these companies to make it more difficult for criminals to use stolen credit cards, while criminals use more sophisticated techniques and/or seek "weak links".

Prior to e-commerce, individuals had to physically present their credit card and sign their name at time of purchase. From the mid-1990s personal identification numbers (PINs) were added. But e-commerce allowed for "card not present" fraud. "Hackers obtained credit card details online, and scammers used social engineering to trick cardholders into revealing their account details. Every security protocol was tightened and improved in response: multi-factor authentication, risk scoring, tokenisation (to reduce the number of times sensitive card details are transmitted), and the management of compromised account systems. Magnetic stripes were replaced by more secure micro-chips that are more difficult to read and duplicate and unlock payments with one-time codes. The latest round of innovation centres on biometric and geolocation verification" (Baker and Shortland 2023 p194).

Baker and Shortland (2023) made a number of observations in conclusion: "First, insurance applies a market logic to crime. The 'economically efficient' level of crime is non-zero: insurers invest in crime control only until the marginal cost of crime reduction equals the marginal benefit to insurers (Coyne and Leeson 2009). Second, insurance and crime can find different equilibria in different times and contexts, ranging from low levels of insurable crime to partially insurable high crime levels. Third, such equilibria are unstable: cost and benefits of either side may shift through exogenous shocks or through endogenous innovations. Fourth, we observed a general pattern of insurers tackling crime at two levels: they help to create significant (and rising) barriers for opportunistic criminals while engaging in strategic interactions with sophisticated and organised crime" (p195).

REFERENCES

Avraham, R & Porat, A (2022) The Dark side of insurance University of Texas Law, Legal Studies Research Paper (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4203765)

Baker, T & Shortland, A (2023) How crime shapes insurance and insurance shapes crime Journal of Legal Studies 15, 1, 183-200

Becker, G (1968) Crime and punishment: An economic approach Journal of Political Economy 76, 2, 169-217

Coyne, C & Lesson, P (2009) Who's to protect cyberspace? In Carr, I (ed) Computer Crime New York: Routledge