UNIVERSITY OF EDINBURGH PHILOSOPHY SOCIETY JOURNAL

Undergraduate philosophy papers published for a general academic audience

Issue 2: 2017-18

List of Editors:

Maia Almeida-Amir

Joshua Cox

Alice Dillon

Jo-Anna Hagen Schuller

Rebecca Stoll

Abi Thwaites

Table of Contents:

A Letter from the Editor-In-Chief:

Since 1871, The Undergraduate Philosophy Society of the University of Edinburgh has provided a space for students to discuss philosophical ideas in ways that go beyond the typical classroom arrangement. In recent years, this has involved inviting world-class lecturers to give a talk every week at our society, as well as organizing events for people to meet and chat about philosophy—with varying degrees of technical formality. During the 2011-2012 academic year, the range of the society's activities were expanded to include an undergraduate journal. Published in the winter of 2012, the first edition comprised three highly interesting papers on decision theory, philosophy of religion, and epistemology. This is the second edition, the first publication since the debut, and I am therefore particularly pleased to have played a role in its publication.

This edition covers a broad range of topics, with articles discussing issues in epistemology, philosophy of religion, philosophy of science, metaphysics, ethics, phenomenology, and existentialism. Much like the last edition, there is no unifying theme — we encouraged members to submit a paper on a topic of their choice. I believe there is a certain value to this approach, given that it reflects the diversity of the interests of our members. It is similarly refreshing to have submissions from authors of different academic backgrounds. While the majority of our papers come from members who study philosophy here at Edinburgh, we also have papers written by students of Computer Science and of Mathematics, as well as a researcher in Chemistry and Medical Sciences.

In addition to all those who submitted papers, I would like to thank the other editors of this journal. Their help was invaluable, and I am very proud of the standard of feedback we were able to provide to our authors during the editing process.

Without further ado, I do hope you enjoy the journal, and I look forward to seeing what direction the next committees take with this project.

Yours Sincerely,

Joshua Cox,
Editor-in-chief

Award for Best Essay & Runner Up

This year's edition introduces a first prize and runner-up award for the best essay. The winning papers this year are:

1st prize: *The Rejection of Normative Supervenience* by Ben Jenkins

This essay is devoted to a problem of supervenience that faces some forms of non-naturalist realism. In meta-ethics, non-naturalist realism is the view that ethical claims reflect real phenomena; but these are phenomena that are not of the same kind as that which we discover scientifically, nor can they be reduced to scientific or natural phenomena. Thus, they are *non-naturalist*. If this is the case, then, there must be some relationship between the two phenomena. Often, what as known as a supervenience relation is posited to explain the relationship between the two. Supervenience in this instance argues that a change in ethical judgement cannot occur without the change of some fact about the natural world. However, meta-ethicists have often taken issue with the explanation that supervenience brings to the table: why is it that the non-natural should supervene on the natural, especially if we cannot reduce the former to the latter? The author looks at one solution to this problem, the complete denial of supervenience within non-naturalist realism, but argues that this strategy is not viable for non-naturalists.

The essay not only picks out a topic of appropriate size, but furthermore grapples with it in a compelling way, both by looking at up-to-date literature, and by developing a novel and convincing response.

Runner up: *Does the distinction of temporal parts effectively differentiate the endurantist and perdurantist approaches to persistence?* by Violet Tinnion

This paper addresses a question regarding personal identity: it argues that the distinction of temporal parts is not sufficient to fully differentiate two views about how objects and individuals persist through time. Under the account the author criticizes, the main difference is that endurantists do not believe that objects have temporal parts, whereas perdurantists do. That is, if we were to trace a line through my journey as an entity through space and time, on the endurantist account, we would say that I exist fully and without division of parts at any moment on this line. Perduantists, on the

other hand, argue that our 'full selves' are made up of lots of distinct temporal parts, and as such we exist only partly at any one time.

Similarly, the author here has done a good job of finding an issue that is not too broad for a short paper, and of exploring it thoroughly while developing a convincing position.

Robustness Analysis: An Application of Weisberg's Toolkit to the Phillips Curve Macroeconomic Model

Robert Campbell

"The nature of robustness analysis is best appreciated through an examination of real scientific examples, not abstract philosophical analysis." – Michael Weisberg (2006a, 734)

I. Introduction

The notion of robustness in scientific modeling was first introduced in Richard Levins' influential 1966 paper, "The Strategy of Model Building In Population Biology". In it, Levins argues that one can discover substantive scientific principles by examining "robust theorems", which occur when differently idealized models of the same target system make similar predictions (Levins 1966). There have been several philosophical attempts to explicate and refine Levin's notion of robustness, most notably those of Steven Orzack and Elliot Sober (Orzack & Sober 1993), William Wimsatt (Wimsatt 1981), and Michael Weisberg (Weisberg 2006a, 2006b). This paper juxtaposes Weisberg's account of modeling to that of Orzack and Sober, and offers evidence in support of Weisberg's account of how robustness analysis is used to assess model predictions. In sections II and III, I present Sober and Orzack and Weisberg's account of robustness analysis and consider three possible objections to Weisberg's account. In parts IV and V, I present the criticisms of the phillips curve model of inflation and unemployment in the late 1960s as an example of Weisberg's account of robustness analysis in action.

II. Two Accounts of Robustness

Under Orzack and Sober's interpretation of Levins, 'robust theorems' are hypotheses about a target system (i.e. system being modeled) whose truth is predicted by all available models of the system in question. Using this framework, they demonstrate a special case under which the robustness of a theorem is indicative of its truth:

> Suppose we know that one of a set of models $M_1, M_2, \ldots, M_n$ is true, but we do not know which. If R is a robust theorem with respect to this set, then R must be true. That is, the following argument is deductively valid:
>
> M_1 or M_2 or . . . or M_n is true.
>
> For each i, M_i implies that R is true.
>
> ---
>
> R is true.

(Orzack and Sober 1993, 583).

Orzack and Sober rightly suggest that this special case is an uncommon one. Rarely does a modeler ever know that a model is True, and rarer still does she know that an unspecified model within a countable set of models is True. In the more typical case where a modeler faces a collection of models that all explicitly contain falsifying idealizations, Orzack and Sober point out that robustness analysis as they construe it cannot be used to establish the truth of a given theorem. Robustness of a theorem with respect to these "false" models does not ensure the theorem's truth (Orzack and Sober 1993, 539). In general, Orzack and Sober claim robustness analysis fails as a confirmation procedure because it is a non-empirical method of confirmation—one cannot confirm the truth of a model merely by looking at other models (Orzack and Sober 1993, 538). Robustness analysis seems to have hit a dead end.

Weisberg claims to offer a more fruitful account of robustness. Under his analysis of Levins, 'robust theorems' are conditional statements that connect a core 'common' structure shared by multiple models to a shared prediction or 'robust property' generated by those models. Robust theorems then take the form "Ceteris paribus, if [common causal structure] obtains, then [robust property] will obtain" (Weisberg 2006a, 738). An example given by Weisberg: the Volterra principle is a robust theorem which states that for ecosystems where the size of a predator population depends on the size of a prey population, a general biocide will increase the size of the prey population relative to the predator population (Weisberg 2008). Here the dependence of the predator population on the prey population is the *common structure*, which gives rise to the *robust property* that a general biocide will increase the share of prey.

Under Weisberg's account, 'robustness analysis' is the process a modeler performs to identify robust theorems, which Weisberg argues consists of the following four steps:

1. A theorist discovers a set of models of a target system that exhibit a robust property.
2. The theorist examines the models in question to reveal why they exhibit the robust property. Here it is assumed that that the models in question make shared predictions by virtue of a *common structure*, which produces the robust property[1].
3. The theorist assigns the parts of the (abstract) common structure of the model to parts of the (real or imagined) target system. She may then substitute the findings from parts one and two into the sentence 'Ceteris paribus, if [common causal structure] obtains, then [robust property] will obtain' (Weisberg 2006, 738). This is the robust theorem.

[1] I suggest in section III that this assumption is in need of further support.

4. The theorist attempts to weaken or remove the ceteris paribus clause by tweaking elements of the models in question to determine under what conditions the robust theorem remains true.

If robust theorems are merely conditional statements linking structures to properties, it is unclear what role they can play in making predictions about actual target systems. This appears to be the crux of Orzack and Sober's criticism—the jump from conditional claims about abstract systems to unconditional claims about actual target systems seems somehow illegitimate.

Weisberg (2006) attempts to build a bridge between the hypothetical and the actual using a technique called 'low-level confirmation'. On my reading of Weisberg, 'low-level confirmation' amounts to answering the question, 'if the target system is structured the way we think it is, are the model's we're using *in principle* capable of representing it?'. The 'in principle' clause is important here—the claim is not that a modeler performs low-level confirmation by regressing model predictions against target system observations to determine whether the model is empirically accurate. Instead, the modeler asks herself whether her model veridically represents the sorts of causal relationships she already thinks obtain in the target system. For example, an economic modeler might ask, 'given that I think an increase in the price of a good will reduce the quantity demanded, does my demand curve model capture this relationship?' The answer to this question constitutes low-level confirmation.

Low-level confirmation is supposed to convince the modeler that the structure of the target system under examination is accurately represented by the common structure of the models that gives rise to the robust property. Having been convinced of this, she may predict that the robust property, as a consequence of the common structure, will appear in the target system. To summarize in a step-by-step format, the full story of how robustness analysis is used in generating predictions looks like this:

1. Low-level confirmation tells a modeler that the structure inherent in some set of her models is in principle capable of representing some target system.
2. Robustness analysis tells the modeler that a robust property of these models is attributable to the common structure that she confirmed using low-level confirmation, rather than the simplifying assumptions of any given model.
3. Now convinced that that the common structure produces the robust property, and that the common structure occurs in the target system, the modeler may predict that the robust property will appear in the target system.

III. Objections to Weisberg's Account

This section addresses three potential objections to Weisberg's account of robustness analysis, two of which I think are not persuasive, and one of which requires additional work to address.

The first potential objection involves low-level confirmation: one could ask whether low-level confirmation is a useful tool in cases where modelers have incorrect assumptions about the casual relationships at work in a target system. It is no use, one might say, to produce models that fit one's intuitions about a target system if these intuitions are incorrect.

My inclination here is to accept the letter of the objection but not the spirit. Yes, it is clearly not useful to produce intuitively attractive but empirically unsuccessful models of target systems. That said, I think this objection only serves to explain why robustness analysis appears where it does in our scientific landscape. Let's say that we accept low-level confirmation as a critical part of making robustness analysis useful, and that low-level confirmation is only effective in cases where we have an accurate intuitive grasp of the causal forces at play in our target system. Given this, it is unsurprising that robustness analysis is popular in population biology and economics, two disciplines where the basic causal forces at play are often quite intuitive and infrequently subject to change. By this I mean that basic assumptions such as 'people make consumption choices (in part) based on prices' are both highly intuitive and seemingly unlikely to be revised. These two features make low-level confirmation more likely to yield correct results and so contributes to the usefulness of robustness analysis.

The first objection also gives us the tools to explain where robustness analysis *doesn't* appear. Because robustness analysis depends indirectly on the basic causal forces in the target system being intuitive and uncontroversial, one would expect robustness analysis *not* to appear in disciplines where the basic causal forces at play are unintuitive or under dispute. This may explain why robustness analysis doesn't appear in disciplines such as physics, where the basic forces are quite unintuitive, and have undergone much revision in the past 100 years with seemingly more to come if current contradictions are to be resolved.

In summary, the objection that low-level confirmation only works when we have accurate intuitions about the basic causal forces at work in our target system places a reasonable constraint on where robustness analysis can be fruitfully applied. However, this constraint does not preclude the use of robustness analysis in situations where the basic forces at work in a target system are intuitive and uncontroversial.

The second potential objection is related to the first objection. Even if we think that our intuitions are a good guide to the causal forces at work in some target system, we cannot know with *certainty* that we aren't mistaken. Why then would modelers ever employ a technique like low-level confirmation, which for this reason is always uncertain?

Again, I'm inclined to accept the letter but not the spirit of this objection. Yes, low-level confirmation can clearly be used to support inaccurate models that ultimately result in bad predictions. However, low-level confirmation still may be the

best procedure for improving predictions available to modelers operating under epistemic constraints. As the example in the next section indicates, low-level confirmation can be particularly useful in cases where data on the relevant phenomenon are scarce, muddy, or otherwise insufficient to immediately confirm or disconfirm a model prediction empirically.

Like the previous objection, this objection gives us a hint for where to find low-level confirmation—it is more likely to appear in disciplines where there is a shortage of experimental data. Sadly, this lacuna appears in both population biology and macroeconomics, where it is frequently too costly to human and animal welfare to conduct experiments that might yield definitive results. In such disciplines, it is unsurprising that modelers turn to alternative techniques to test their models that rely less heavily on experimental data. My example in the coming section is chosen to indicate how robustness analysis and low-level confirmation can be used to replace bad models with better ones, even in the absence of the relevant falsifying data.

The third possible objection takes offense with step two of Weisberg's account of robustness analysis, where the modeler examines the models that give rise to the robust property in order to discover the common structure. Weisberg assumes that diverse models which give rise to a robust property will do so by virtue of a common structure, presumably a structure also instantiated in the target system. Weisberg implicitly uses this assumption to explain how the robust property appears in both the models and the target system; in both cases, the common structure creates the robust property seen in the model and target system.

This picture implies that empirically successful models of a target system must veridically represent that target system, by which I mean that relevant structural properties of the target system must appear in the successful model. However, some analyses of models, including that of Milton Friedman (Friedman 1953), suggest that non-representational models with high unrealistic structures may still make consistently successful predictions. For a detailed account of Friedman's position and the argument for modeling without representation, see (Isaac 2012).

The possibility of modeling without representation seems to present a problem for Weisberg's account of robustness analysis. If it is possible for multiple models with non-representational and varied structures to independently and correctly predict the behavior of a target system, then there would appear to be many robust properties that do not correspond to common structures. This is a problem for Weisberg because it implies that the common structures, which are central to his account of modelling, are less important than previously thought. If common structures are actually non-essential to robustness, this would cast doubt on Weisberg's account. For this reason, Weisberg's tacit assumption that successful models accurate represent the structure of their target system is in need of support.

Section IV. A Tale of Two Models.

This section tells a story about the 'phillips curve' model of inflation and unemployment, which suffered a large loss of faith within the field of economics in the late 1960s. At the time that the phillips curve was being discredited, empirical support for its predictions were at an all time high, and the first wave of disconfirming data was still several years away. This raises a question for philosophers of science: why did economists lose faith in the phillips curve *before* its predictions began to fail? In section V, I will attempt to answer this question using the conceptual tools from Weisberg's account of robustness. On to the story…

In 1958, William Phillips published a now famous paper documenting the apparent relationship between inflation and unemployment in the UK (Phillips 1958). In it, he noted how a regression of inflation data against unemployment data for the UK from 1861 to 1913 appeared to show a nonlinear, inverse relationship between inflation and unemployment.

Fig 1. (Phillips 1958, 285)

This relationship was developed into the 'phillips curve' model of inflation and unemployment, most notably in a 1960 paper by Paul Samuelson and Robert Solow, which presented the phillips curve as "the menu of choice between different degrees of unemployment and price stability" (Solow & Samuelson 1960, 192). There were several other early variations on the model, all of which robustly predicted that governments could reduce unemployment in exchange for inflation in both the short

and long run. I will lump all of these versions together under the term 'the basic phillips curve'.

The causal story behind the basic phillips curve was this: if the government attempts to increase aggregate demand[2], it will induce firms to produce more goods in order to satisfy the increase in demand. In order to produce these goods, the firms will have to hire more workers, many of which will come from the pool of unemployed workers. As firms compete to scrape the bottom of the unemployment barrel, they will raise their wages in order to gain more workers and hold on to the workers they already have. These increased wages will cause inflation. In short, a country can have permanently low unemployment, so long as it is willing to stomach permanently high inflation (Gottfries 2008, 242).

The initial statistical support for the basic phillips curve was strong. Central banks in many countries examined their historical data on inflation and unemployment and found statistical evidence of a tradeoff between inflation and unemployment (Mankiw & Reis 2018). By the mid 1960s, the basic philips curve was unquestioningly accepted by the field of macroeconomics, and had become a staple of period monetary policy (Hall & Seargent 2018). Many countries, including the US, began to pursue inflationary monetary policy in the hopes of reducing unemployment. In 1961, the Federal Reserve raised the growth rate of the M3 monetary supply from a modest 4% to what would become a decade long average of 7% - 8%, a policy guaranteed to create inflation (Fred.stlous.org 2018).

The confidence in the long run tradeoff between inflation and unemployment was however short lived. In 1968, the economist Milton Friedman gave a Presidential address to the American Economic Association where he proposed a model of unemployment called the 'natural rate of unemployment' (Friedman 1968). In direct contradiction to the phillips curve, the natural rate of unemployment model predicted that unemployment would trend towards an equilibrium value in the long run irrespective of monetary policy.

The causal story behind the natural rate of unemployment model is this: if inflation is consistently high, firms and workers will come to expect high inflation. If everyone expects high inflation, then workers will negotiate for rapidly rising wages in order to preserve the purchasing power of their income. If wages are rising in line with expected inflation, then an expected inflationary increase in goods prices will not induce firms to hire more workers, since the extra money firms could earn by producing more units of the good is offset by the extra cost of hiring more workers at the new higher wage to produce those units. In short, one cannot use inflation to decrease unemployment below the 'natural level' in the long run, since people will come to expect inflation and adjust accordingly (Mankiw & Reis 2018, 84). As Friedman put it in his presidential address, "there is always a temporary tradeoff

[2] By lowering interest rates or increasing the monetary supply

between inflation and unemployment; there is no permanent tradeoff"(Friedman 1968, p11).

When Friedman gave his presidential address in 1968, the basic phillips curve enjoyed very strong empirical support (Mankiw & Reis 2018, 83). Western central banks had seen marked success at reducing unemployment using inflation in the short and medium run during the 1960s, and the Federal Reserve had recently completed a highly sophisticated and (temporarily) accurate macroeconomic model called the MPS that employed insights from the basic Phillips curve (Brayton and Mauskopf 1985). At the same time, early analyses of the natural rate of unemployment model seemed to indicate that Friedman was wrong, or at least overestimating the importance of inflation expectations in actual realized inflation (Hall and Sargent 2018, 132)[3]. In fact, the available data so clearly favoured the phillips curve that Friedman's prediction of 1970s "stagflation", which violated the phillips curve by exhibiting both high inflation and high unemployment simultaneously, is still regarded as, "one of the greatest successes of out-of-sample forecasting by a macroeconomist" (Mankiw and Reis 2018, 88).

Given the preponderance of evidence in 1968, one would predict that Friedman's criticisms of the phillips curve would have been laughed out of the room, or at least put on hold until the data supporting his long run predictions came in during the 1970s. This did not happen. There is very strong evidence to suggest that faith in the long run tradeoff between inflation and unemployment began to crumble almost immediately after Friedman's presidential address, almost five years before blooming 1970s stagflation brought "wreckage" to the basic phillips curve (Gordon 2008). In 1968, Robert Solow[4], the original populariser of the phillips curve, began including an inflation expectations parameter in his model of inflation, admitting by this fact that he didn't believe that the basic phillips curve held in the long run (Solow 1968). James Tobin[5] did the same (James 1968). Robert Gordon, an economist famous for his work on growth forecasting, soon followed suit (Gordon, Solow & Perry 1970). Even the Federal Reserve appeared to have lost some faith in expansionary monetary policy in 1970, as it dropped the growth rate of the M3 monetary supply from a decade long average of 7% - 8% to 4%, before being forced raise it again due to recession (Fred.stlouis.org 2018).

Though the long run tradeoff between inflation and unemployment was not completely discredited until the global recession of 1973 – 1975 provided undeniable evidence of stagflation, there was clearly an early loss of faith that significantly predated the falsifying data. This should create some *prima facie* confusion for philosophers of science; why did professional economists discount long run predictions by the basic phillips curve years before they were shown to be inaccurate?

[3] These analyses were later shown to systematically underestimate the importance of expectations.
[4] Nobel Prize 1987
[5] Nobel Prize 1981

In the coming section, I argue that one can answer this question using Weisberg's account of robustness analysis and low-level confirmation.

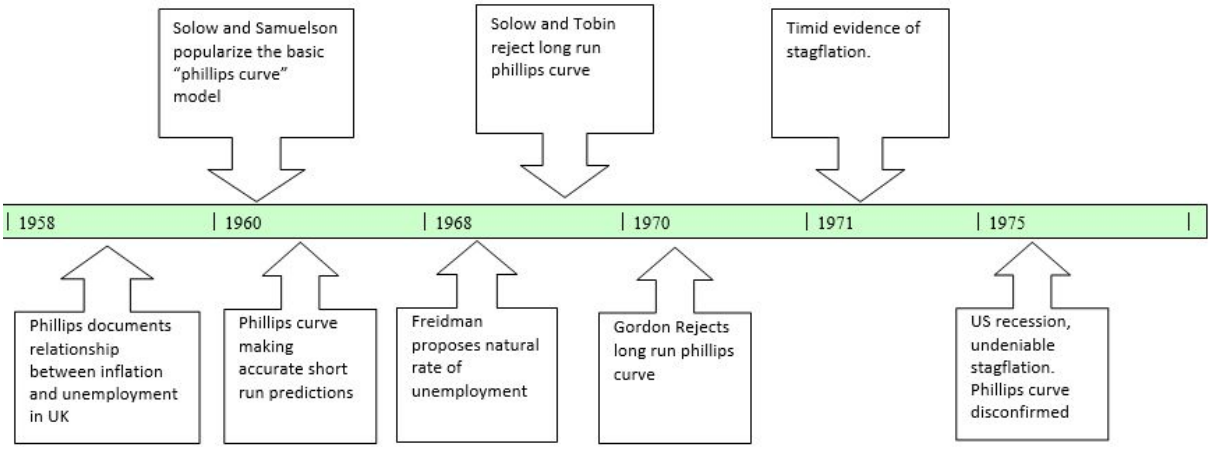
Section V. Analysis of the Tale of Two Models

The parable of the phillips curve and the natural rate of unemployment makes real the idea that modelers are sometimes forced to assess the verticality of models in situations where the relevant data is unavailable. Surely, a successful account of modeling must be able to explain how it is that a modeler may do this. This section will give an account of why economists began discounting the tradeoff between inflation and unemployment predicted by the phillips curve using the conceptual tools found in Weisberg's account of robustness.

Weisberg's Toolkit glossary

1. *Robust properties* are predictions shared by multiple models of a given target system.
2. *Robustness analysis* is a process of discovering the common structure that leads models to exhibit a robust property.
3. A *robust theorem* is a conditional claim linking a common structure to a robust property.
4. *Low-level confirmation* is a technique for examining whether a model is in principle capable of representing a target system.

Phillips Curve Timeline



In the eight years between the popularization of the basic phillips curve and Friedman's proposal of the natural rate of unemployment, the tradeoff between inflation and unemployment was a robust property of all available models of inflation and unemployment. The natural rate of unemployment predicted that the tradeoff did not hold in the long run, meaning that the tradeoff was no longer a robust

property with respect to any set of models that included the natural rate of unemployment.

As Orzack and Sober pointed out, the fact that a prediction is not robust with respect to all available models is not by itself good reason to doubt it (Orzack & Sober 1993, 68). To know whether disagreements among models should be taken seriously, one must first know *why* the models in question disagree—importantly, whether they disagree because of a difference in simplifying assumptions or because of a difference in core structure. A modeler can answer this question by performing something like the inverse of step two in Weisberg's account of robustness analysis. Instead of looking for a common structure to explain a robust property of a model, the modeler looks for a difference in structure to explain conflicting properties of two models.

In the case of basic phillips curve and the natural rate of unemployment, a modeler could have discovered the source of the disagreement by examining the mathematical properties of the two models. The basic phillips curve claimed that inflation (π) in a given period was a function of unemployment (u) and unspecified price shocks (z), represented by the following functional form:

$$\pi = f(u, z)$$

A simple modern inflation model (Gottfries 2013)[6] that accounts for the natural rate of unemployment predicts that that inflation (π) in a given period is a function of deviation from the natural rate of unemployment ($u - u^n$), expected inflation (π^e), and unspecified price shocks (z), represented by the following functional form:

$\pi = f(u - u^n, \pi^e, z)$, where π^e is an increasing function of π in previous periods.

We can see that these models differ with respect to their core structure by examining the different variables that appear in their functional forms. The modern inflation model predicts that inflation expectations (π^e) and deviation from the natural rate of unemployment ($u - u^n$) influences inflation, while the basic phillips curve does not include a measure of inflation expectations and does not posit the existence of a natural rate of unemployment. These models differ with respect to their core structure in part because they posit very different sorts of entities to explain inflation.

Further analysis of the models' core structures reveals why they make different long run predictions about inflation and unemployment. The functional form of the basic phillips curve claims that inflation is a simple function of unemployment, which indicates that there is a stable relationship between the two. This creates the possibility for a long run tradeoff. Conversely, the functional form of

[6] Modern inflation models that include unemployment are actually still called philips curves. They are distinct from the 'basic phillips curve' described in this paper in that they take into account the natural rate of unemployment. I will violate convention and call them 'inflation models' for the sake of clarity.

the modern model claims that inflation is in part a function of inflation expectations, and that inflation expectations are determined by an increasing function of inflation in previous periods. This means that an extended period of high inflation will eventually cause inflation expectations to rise, which will cause a further (permanent) increase in inflation without changes to unemployment. In other words, the structure of the modern model implies that there is no stable tradeoff between inflation and unemployment.

So far we've determined that our two models make different predictions about the tradeoff between inflation and unemployment, and that these different predictions are caused by their different core structures. In order to determine which of these model predictions is correct, a modeler must then determine which model's structure (if either) is instantiated in the target system. If reliable empirical methods for determining this are not readily available, she may attempt to tackle the question using low-level confirmation. In this case, low-level confirmation would require a modeller answer the following question: 'given my understanding of the macroeconomy, does the phillips curve or the natural rate of unemployment model better represent the structure of this system?' Focusing in on the key differences between the two models, the modeller could narrow the question down to: 'given my understanding of the macroeconomy, do I think that inflation expectations play a role in determining inflation?' This last question can be fruitfully debated with very little data, and the answer the modeler comes to will allow her to make a strong prediction about the tradeoff between inflation and unemployment.

The modeler can reach the correct answer to this question using a simple thought experiment that frequently appears in introductory macroeconomics courses. Imagine that current inflation is at 0%, and that the economy is in equilibrium. *Ceteris paribus,* if every firm expects prices to rise by 2% in the following period, they will all individually raise their prices by 2% to stay in trend, meaning that realized inflation across the economy will be 2%. At the very least, it would be incredible if every firm raised prices in response to expected inflation, and the overall price level *did not* inflate. Accordingly, inflation expectations have an impact on realized inflation. This analysis is roughly true and cuts to the causal heart of the issue. Most importantly, it (indirectly) allows the modeler to reach a substantive conclusion about the long run tradeoff between inflation and unemployment given very little data. Below is a step-by-step summary of how the modeler can reach a conclusion about the trade off:

(1) The modeler notices that the phillips curve and natural rate of unemployment make different predictions about the long run trade off between inflation and unemployment. ie, the phillips curve prediction of the long run tradeoff is no longer a robust property.
(2) The modeler performs a kind of robustness analysis whereby she determines whether the disagreement between models is attributable to a difference in

core structure or to a difference in simplifying assumptions. She discovers the different predictions stem from a difference in core structures.

(3) The modeler performs a low-level confirmation on the two core structures to determine which better represents the causal forces believed to be at work in the target system. Low-level confirmation reveals that the structure of the natural rate of unemployment better represents the causal forces at play. The phillips curve is less capable of representing the target system because it ignores the causally potent effects of inflation expectations.

(4) Having determined that the core structure of the natural rate of unemployment better represents the structure of the macroeconomy, and that this structure entails that there is no long run tradeoff between inflation and unemployment, the modeler may infer that the macroeconomy will not exhibit a long run tradeoff between inflation and unemployment.

Several features of the above analysis bear mentioning. Perhaps most importantly, the process of robustness analysis and low-level confirmation described above is far from infallible. In particular, the success of this low-level analysis hinged critically on the modeler having intuitions of sufficient strength and accuracy to engage in fruitful *a priori* contemplation of the target system. For example, a modeler employing the low-level confirmation described above would have to understand what it means to have expectations, and understand that firms set their prices in order to increase profits. Thankfully these concepts are fairly intuitive to most humans. If they were not, the results of this low-level confirmation would be much less certain. This should tell us that robustness analysis and low-level confirmation are most likely to be useful in disciplines that model target systems where the basic causal forces at play are intuitive enough to be fruitfully considered *a priori*. Again, I point to economics and population biology as two such disciplines.

Despite the possibility of error, using robustness analysis and low-level confirmation to check model predictions seems clearly better than not checking predictions at all. In the case of the phillips curve and the natural rate of unemployment, the alternative to this procedure was to wait three years for the basic phillips curve's monetary policy prescriptions to result in recession and stagflation. Given the unattractive nature of this prospect, it is unsurprising that economists opted to employ robustness analysis and low-level confirmation before the data rolled in. This offers some support to my argument in section III that robustness analysis and low-level confirmation are more attractive in cases where immediate empirical refutation or confirmation of a model prediction is unavailable or costly.

Section VI. Conclusion

The first three sections of this paper characterized Weisberg's account of robustness analysis within the context of competing views, and presented three objections and responses to his account. In sections IV and V, I considered a surprising event in this history of macroeconomic modeling, and used Weisberg's conceptual toolkit to explain this happening. The primary takeaways from this

exercise are as follows: (1) robustness analysis and low-level confirmation can play an important role in evaluating model predictions. (2) low-level confirmation is an empirically risky procedure that is most attractive when standard empirical confirmation is not immediately available. (3) low-level confirmation is most likely to be effective in cases where the modelers performing the confirmation have a strong and accurate intuitive grasp of the casual forces at work in the target system.

If correct, these three findings offer hints on where to find more cases where robustness analysis and low-level confirmation are used to assessing predictions. Psychology, for example, seems to be a discipline where the basic causal forces are often intuitively graspable, and where ethical considerations places limitations on the possibility of experimentation. Sociology may also share these features. Further attempts to examine the prevalence of robustness analysis and low-level confirmation may consider examining these two disciplines as well.

References:

Brayton, & Mauskopf. (1985). The federal reserve board MPS quarterly econometric model of the US economy. Economic Modeling, 2(3), 170-292.

Fred.stlouisfed.org. (2018). M3 for the United States. [online] Available at: https://fred.stlouisfed.org/series/MABMM301USA657S [Accessed 15 Feb. 2018].

Friedman, M. (1953). The methodology of positive economics. In *Essays in positive economics*. 3-46. Chicago: University of Chicago Press

Friedman, M. (1968). The Role of Monetary Policy. *The American Economic Review, 58*(1), 1-17.

Gordon, R. J. Solow, R. Perry, G. (1970). "The Recent Acceleration of Inflation and Its Lessons for the Future". Brookings Papers on Economic Activity, 1970(1), 8-47.

Gordon, R. J. (2008). *The History of the Phillips Curve: An American Perspective*. Lecture presented at Australasian Meetings of the Econometric Society: Keynote Address, Auckland New Zealand.

Gottfries, N. (2013). *Macroeconomics*. Hampshire, United Kingdom: Macmillan Publishers Limited.

Hall, R., & Sargent, T. (2018). Short-Run and Long-Run Effects of Milton Friedman's Presidential Address. Journal of Economic Perspectives, 32(1), 121-134.

Isaac, A. M. (2012). Modeling without representation. *Synthese, 190*(16), 1-13.

Levins, R. (1966), "The Strategy Of Model Building In Population Biology," *American Scientist*. 54(4): 421-431

Mankiw, N., & Reis, R. (2018). Friedman's Presidential Address in the Evolution of Macroeconomic Thought. Journal of Economic Perspectives, 32(1), 81-96.

Orzack, S. H., and E. Sober (1993), "A Critical Assessment of Levins's 'The Strategy of Model Building in Population Biology' (1966)," *Quarterly Review of Biology*. 68(4): 533-546.

Phillips, A. (1958). The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–1957. Economica, 25(100), 283-299.

Samuelson, P., & Solow, R. (1960). Analytical Aspects of Anti-Inflation Policy. The American Economic Review, 50(2), 177-194.

Solow, R. M. (1968). *Recent controversies in the theory of inflation*. S. Rousseas (Ed.), Proceedings of a symposium on inflation, Kazanjias Economics Foundation, Wilton, CT.

Tobin, J. (1968). *Discussion*. S. Rousseas (Ed.), Proceedings of a symposium on inflation, Kazanjias Economic Foundation, Wilton, CT.

Weisberg, M. (2006a), "Robustness Analysis," *Philosophy of Science*. 73(5): 730-742

Weisberg, M (2006b). "Forty Years of 'the Strategy': Levins on Model Building and Idealization," *Biology and Philosophy*, 21(5): 623-645

Weisberg, M. and K. Reisman (2008). "The Robust Volterra Principle," *Philosophy of Science*. 75(1): 106-131.

Wimsatt, W. C. (1981), "Robustness, Reliability, and Overdetermination," in M. Brewer and B. Collins (eds.), *Scientific Inquiry and the Social Sciences*. San Francisco: JosseyBass, 124-163.

Understanding emotion phenomenologically

Cezar Mihalcea

It is undeniable that we experience a world with which we interact, and which leaves its mark upon us. In that sense, we are subjects of the world, firsthand, as a self-determined horizon of our being (as what we could be), and in an intersubjective, actual manner (objectively), as reality. Furthermore, we experience ourselves, both as a part of the world and as someone who experiences the world from outside, a transcendental ego (Kant, 1787), at all moments being the same subject of all this experience (unitary in all our individual moments of existence, self-identical). Subsequently, we can argue that emotion is part of our experience, pertaining to the apprehension of the subjective self, as a part of it. However, for us, our emotions appear as more than mere figments of our imagination, they appear as real, almost tangible things. It is for precisely this reason that we must treat emotions not as belonging to the purely transcendental ego, but as authentic phenomena that unfold within our stream of consciousness[7]. For this reason, we will unveil the essence of emotion (within experience) through a phenomenological analysis. It is also noteworthy that, within this analysis, we are considered fundamentally embodied beings. Thus, we, as transcendental subjects of our experience, are identical to our psycho-physical selves (which is how we experience all other humans, as living, thinking beings, but exterior to our own mind), which enables the reconciliation of subjectivity and objectivity in our understanding of phenomena. (Merleau-Ponty, 1945)

In our investigation of emotion, we will build upon the principles established by Edmund Husserl in the Cartesian Meditations. We will begin by explaining the phenomenological reduction and what it entails, and why the representations of phenomena we obtain are essential to the elucidation of experience. Furthermore, we will attempt to describe civilization as a transcendental, intersubjective foundation for objective experience. Lastly, we will establish a phenomenological explanation of emotion, and we will consider to what extent emotion could be objective, or even be found within civilization rather than being isolated within an individual.

The basis for subjective experience. The transcendental reduction and its structures.

We begin by explaining the structure of experience phenomenologically, by means of a *phenomenological reduction* (Husserl, 1950). This attitude to phenomena

[7] Here, we do not imply that emotions cannot be unconscious. Rather, the stream of consciousness, phenomenologically speaking, is the context of all experience, both conscious and unconscious, real or possible. This rightly serves our purpose of elucidating emotion as a phenomenon to which we are subject, experiencing it constantly. It follows that we will not describe the ego in its Freudian sense, but rather in its phenomenological sense.

is the reduction of my experience of the objective world and self to the realm of my unitary (self-identical) *transcendental ego* (which is apodictic to experience - "I am" precedes what I perceive). This should clarify the internal structure of thought and knowledge; the world is now nothing more than what I perceive and think. It does not "disappear", but it only temporarily gains the property of lacking a fundament for the claim of objective existence, becoming actually evident only for me. The idea of a transcendental ego had been explored by Kant (1787) before, corresponding to the ideal of a unitary subject[8], but with Husserl we are faced with an extended concept of the ego, for which self-reflection is possible. Specifically, by reducing all possible experience to the transcendental ego, I do not lose grasp on reality, but instead reality becomes ever-present within my perception as a phenomenon. Apperception (transcendentally-reduced perception) becomes all-encompassing of phenomena, whereas for Kant, the phenomenon is exterior to apperception. With Husserl, my living self in the world is identical to its representation within my transcendental ego - I can perceive my actual, real self.

Moreover, in the Second Meditation Husserl introduces the concept of intentionality, as a universal structure of possibility for any given phenomenon. Experience becomes the appropriation of apperceived phenomena, which Husserl calls *evidence*, to the corresponding intentionality, which is a horizon of possibility. When it comes to real (evident) objects, this appropriation is nothing more than a synthesis, developing within the internal temporality of apperception: the object, appearing to me as continuous in its unity (Husserl calls this *the passive genesis*) represents the foundation of the actual, continuous act of constituting the object within my thought (*the active genesis*). The intention is "filled" by the perceived object. The possible forms of the object are suppressed within its sphere, as the actual form becomes evident, though never *fully* evident, as that would require all its traits to be simultaneously perceivable – for instance, I'd need to see both the front and the back of a door. Consequently, the object becomes part of my flux of conscience, as an *object that exists for me*, and an object to which I can permanently return to re-actualize its representation.

For example, let us imagine an apple. From a given angle, I can see what I perceive as the front. The back has an indefinite number of possible shapes. As I rotate the apple, the before-hidden sides are revealed, therefore further reducing the number of possible forms the apple could have. However, the already-seen sides become obscured, which opens them to possibility again, although I already have previous knowledge of their past form. If I were to return to those, and see that someone has taken a bite there, the new knowledge would replace the former one. However, I can never have full certainty of how the apple is, but I do have a temporally determined synthesis of its already-perceived forms. Similarly, internal objects (those only thought, imagined) unfold within a temporal synthesis, but as

[8] The transcendental ego as a unitary subject, in Kant's Transcendental Dialectic, can be considered purely transcendental, therefore lacking substantiality. It is precisely for this reason that, in Kantian philosophy, one cannot have any intuition (and, therefore, knowledge) of oneself.

they have never been brought to evidence, they remain within the realm of possibility.

Finally, and most importantly, Husserl's phenomenological attitude makes self-experience and, with it, the experience of emotion as a phenomenon possible. Since the synthesis of the phenomena (both possible and real) makes up my stream of conscience, I can think of its unity as a phenomenon by itself – my whole thought turned into an object of experience. This naturally leads to the division of my ego into its counterparts: the ego as my stream of consciousness (the *flowing cogito*), through which and inside which all other phenomena unfold, and the ego as the subject of the experiencing of self as described above – Husserl's *ego pole*, the *I* in *I think*. However, I am ever-conscious of my existence as both myself and for myself – the two counterparts are, in fact, unitary, existing within a dialectical identity, which Husserl calls (in the fourth meditation) the *monadic ego*. It is here that emotion as a purely subjective part of my ego finds its identity with itself as a phenomenon.

Objectivity and civilization

In the absence of any actual, perceived phenomena, it is evident that the above described structure of the ego is empty. It is a self-identical, apodictic form that stands as the foundation for the possibility of all experience. In the fifth meditation, Husserl argues for the universality of the monadic ego as the *eidetic* (essential) form of all possible subjects[9], which means that my experience of the world is not purely subjective, but rather intersubjective, at least as far as the phenomenal world that I have experienced is also experienced by the others around. This eidetic form, obtained from my thinking of myself as the the possible form of any other thinking being, serves as a different principle of intentionality, with its corresponding transcendental reduction. This enables my apprehension of any other human being as alterity, a different form of what I could have been, not as an object, but as a psychophysical reality. In other words, any other thinking being is experienced as another monadic ego, possessing both a perceivable actual 'body' and an unperceivable conscience, of which we only know the form. Husserl calls this appropriation of myself to the Other *empathy*. It happens as a form of successive pairing between me and the members of any given group in regards to the corresponding intentionality. He described as follows:

I understand my transcendental ego as an empty horizon for the possibility of thinking; it is, essentially, the intentionality corresponding to the phenomenon of my actual self. In light of this, any such "self" could serve as an actuality for this intentionality. Furthermore, the Other is given to me as his own "self" within the world, by means of the same intentional horizon described above. It is from this that we understand the structure of the Other as "another me", given that I perceive it as both an ontologically different subject of the phenomenal world (an exterior, ego-pole) but also as embedded in it in the same way I am. It follows that all

[9] With this, Husserl attempts to counter the solipsistic view his philosophy would otherwise lead to.

experience bears significance not only within my apperception, but within the apperception of others too, in exactly the same way. Furthermore, on this basis, we can deduce that the world is identically experienced by different subjects (who are essentially structurally the same, considering the common sphere of intentionality), taking shape within an *inter-monadic community*, as an unspoken transcendental convention. In other words, the ego-intentionality determines the common shape of thinking, and, in conjunction with object-intentionality, it constitutes the world as an intersubjective convention. In addition, the evidence that fills object-intentionality leads to the foundation of a primordial stratum of objectivity – the consistent world of real, substantial objects.

However, this is only the first step in establishing intersubjectivity. Besides our experience of objects that we perceive as real, we also experience *meaning*, both by itself (as belonging to us, in our self-experience) and as the meaning of an objective phenomenon (for instance, the meaning of a piece of art). At a first glance, one could argue that this meaning is purely subjective. However, as a subject in the world surrounded by others, I also experience their view of a meaning – as long as it is in regard to an object (since, with no actual real phenomenon to refer to, there would be nothing exterior, common to experience) and they communicate it. I am, of course, influenced to some degree by the views of others, and it is my choice if I disregard them or agree to them. This interaction between me and the others, upon the primordial stratum, functions either towards strengthening a *convention of meaning*, or towards challenging it and establishing a new one – this convention (of a transcendental nature) being precisely what we call a "culture", or, historically speaking, *civilization*, as a secondary stratum of intersubjectivity.

Let us examine more attentively how this construction of meaning is phenomenological. In *The Origin of the Work of Art*, Heidegger (1960) argues that meaning functions similarly to our apperception of the spatio-temporal dimension of any object[10]. Each individual's understanding of the meaning of an object represents a temporally-determined, singular meaning of its being. As one would bring more subjects together, their views would either conflict or converge, in such a way adding more singular meanings, which in themselves form a synthesis too. However, the mechanism behind this synthesis is not temporality (as internally determined, like in the case of objects), but *history* (time in relation to civilization). In other words, as history advances, meaning changes from person to person, and from cultural sphere to cultural sphere (this sphere being a subject-constructed horizon of meaning-possibility) – however, the previously experienced meanings are not lost, but rather fall into possibility again (and into previous history), so that they might re-emerge at some point, given the right circumstances (the *concealment* and *unconcealment* of being are equivalent to the *actual* and *possible* in intentionally-directed apperception). This historical synthesis of meaning is what we call civilization, possibly being the main means through which we have built conventions, rules, morality, etc. It is also necessary to note that the meaning built in

[10] Heidegger also argues for the fact that meaning is built directly on the object's simple existence, and that this simple existence serves as a grounding horizon of possibility for the meaning.

this way is in no way actually existent, but it is abstractly contained in the relation between the object and the subjective experience of it, be it individual or cultural. This is the degree to which we can argue for "objective" ideas, or ideas of culture, and it will serve as the basis for the intersubjectively-experienced emotion.

Emotion

Now that we have set the ground for our phenomenological analysis, it is time that we turn to emotion. We must begin by describing the character of emotion as a phenomenon, subject to intentionality.

Firstly, we can evidently say that, at any point in time, I experience emotion to some degree – whether it is anger, happiness or boredom. I also experience these emotions for a duration of time, and general shifts in emotion are perceived as gradual. It follows that emotion is experienced temporally in much the same way that an object is; any given, particular emotion takes the place of the previously experienced one, which falls back into further possibility, a particular *intentional horizon of emotion*. Also, each separate emotion is impossible to delimit from the other, as I always perceive the change to be gradual.

Yet, similarly to meaning, emotion cannot be experienced by itself, as it possesses no perceivable form; it can't have a spatio-temporal representation. Therefore, there is need for an object, either real or thought of (remembered, imagined), for emotion to be experienced within the subject-object relation, much like meaning. If the object is imaginary, the emotion lies purely inside my transcendental, imagined world, and therefore its only reality rests upon the possibility of the imagined object to exists; it is not actual, evident. However, if the phenomenon in question is real, we can claim the same degree of intersubjectivity as for meaning. Namely, I feel happy in relation to a gift I have received, or frustrated with some arbitrary event that disrupted my work at some time (since events are, essentially, interactions between multiple objects and sometime subjects too, so they can be considered phenomena as well). Also, other people might experience the same – or a different – emotion in regards to the same phenomenon, which results in a similar intersubjective stratum of emotion. Very much as it is with meaning, a civilization can experience, within its cultural sphere, emotion pertaining to an object contained in that sphere[11].

So far, emotion seems to be just another form of meaning, since it unfolds within my egologic world on the basis of other phenomena. Nevertheless, emotion does possess a distinguishing feature: the fact that, in all my apperception of it, I experience it always in relation to my transcendental self. This means that, as opposed to meaning, it always bears an influence upon me, as a part of my ego. This influence is reflected unto the world and unto other subjects of it. For example, I understand the meaning of a certain object, but I am able to keep said meaning to myself – I would have to express it in language for its insertion into culture.

[11] This adds a degree of justifiability to the epistemological basis of judgements such as "The Ancient Greeks felt good about art in general, or about this or that temple".

Conversely, whenever I experience emotion, it has an effect on me, and it is perceivable by other subjects to a certain degree (at least through body-language, in virtue of my permanent embodiment). Also, given that emotion is virtually part of me, I am able to read and experience another subject's emotion too, if only as intentionally-directed possibility, constrained under the evident context. This ever-continuous experience of emotion would suggest that its essence lies in both the subject-object (and culture-object) interaction and the subject-subject (subject-self) interaction, being subjective *and* objective, but, most importantly, omnipresent within my transcendental ego as a phenomenon. Consequently, it is the only consistent way in which we can attain a degree of insight into the flux of consciousness of another person, and truly experience what they experience.

In conclusion, our investigation has led us to an understanding that emotion is not only a subjective phenomenon that one experiences, but also an objective one, unfolding gradually within civilization in relation to real phenomena. Moreover, we have also identified it as a means to mediate a relation between the content of the egologic life of two different subjects, as it always relates to the subject in its perceived existence. Therefore, we have limited access to the actual experienced phenomena of another subject, even though it is immanent to his ego, these phenomena being precisely his emotions. That is why we can safely, although metaphorically, call it a "gateway into the soul", in its uniqueness as a phenomenon.

References:

- Immanuel Kant, 2007. On the Deduction of the Pure Concepts of the Understanding. In *The Critique of Pure Reason* (pp. 124 – 139). London: Penguin Classics
- Maurice Merleau-Ponty, 2012. Preface. In *Phenomenology of Perception* (pp. XX – XXXV). Oxford: Routledge
- Edmund Husserl, 1999. First Meditation: The way to the transcendental ego. In *Cartesian Meditations* (pp. 11 - 26). Dordrecht, The Netherlands: Kluwer Academic Publishers
- Immanuel Kant, 2007. Of the Paralogisms of Pure Reason. In *The Critique of Pure Reason* (pp. 316 - 378). London: Penguin Classics
- Edmund Husserl, 1999. Second Meditation: The field of transcendental experience laid open in respect to its universal structures. In *Cartesian Meditations* (pp. 27 - 53). Dordrecht, The Netherlands: Kluwer Academic Publishers
- Edmund Husserl, 1999. Fourth Meditation: Development of the constitutional problems pertaining to the transcendental ego himself. In *Cartesian*

Meditations (pp. 65 - 81). Dordrecht, The Netherlands: Kluwer Academic Publishers

- Edmund Husserl, 1999. Fifth Meditation: Uncovering of the sphere of transcendental being as monadological intersubjectivity. In *Cartesian Meditations* (pp. 89 – 141). Dordrecht, The Netherlands: Kluwer Academic Publishers
- Martin Heidegger, 1960. The Origin of The Work of Art. Retrieved from http://users.clas.ufl.edu/burt/filmphilology/heideggerworkofart.pdf

A Modern Scientific Philosophical Approach to Religion and its Central Beliefs.

Dr. Christopher N Haggarty-Weir (1-3) and Mr. Bryce Harper (4).

Both authors contributed to this paper. Corresponding author CNHW (email: c.n.weir@sms.ed.ac.uk).

Author Affiliations:

EaStCHEM, School of Chemistry, University of Edinburgh, Edinburgh, United Kingdom.

Walter and Eliza Hall Institute, Parkville, Victoria, 3052, Australia.

Department of Medical Biology, University of Melbourne, Parkville, Victoria 3052, Australia.

School of Communication and Arts, University of Queensland, St. Lucia, Queensland, 4072, Australia.

Introduction

It is hard to imagine humanity without religion and science. Both have coloured the history of civilisation. Whilst the tools and discoveries of science have been used to execute terrible wars, so too has religious sectarianism caused its share of destruction. Richard Feynman, the Nobel Prize-winning physicist and polymath, saw no inconsistency between religion and science; in his eyes you could be both a person of science and faith (Feynman *et al.*, 1999 p. 247). However, science has cast light on specific faith claims of religion, which often form underlying principles and concepts in a given religion. The laws of the natural world – as revealed by the scientific method – present some troubling questions for the validity of divine creation, miracles, and surviving death in the afterlife. This essay will examine those three specific religious claims and will argue that it is science, not religion, that reveals the natural world for how it really is. For conciseness, the focus will be on the Abrahamic religions; though, the underlying philosophical thought and use of the scientific method can be applied ubiquitously to any world religion. First however, this paper will give some background into the perceptions of science and religion.

Perceptions of Science and Religion

Scientific inquiry has been a powerful tool in human progression. As an example, the now famous discoveries of Darwin and Newton illustrate the immense progress that can be made through the scientific method (Ayala, 2009). Often science challenges our intuitions, beliefs, and our orthodoxies. Religious explanations for natural phenomena have given way to scientific ones. Darwin's theory, which takes place over millions of years, seems at odds with the literal 6-day creation of the Bible. Despite new discoveries that seem at odds with religious teaching, religion has

managed to co-exist with science and its expanding catalogue of discoveries. Professor of theology, John F. Haught, has described four different possible relationships between religion and science. It is for this reason that Haught's position will be examined first so his framework can be used for later discussion. Haught (1995 p. 9) claims that religion and science can relate in the following four ways:

1. Conflict – Religion and science are not compatible.
2. Contrast – Religion and science are non-overlapping fields; therefore, there is no conflict.
3. Contact – Religion and science are capable of dialogue and perhaps even harmony.
4. Confirmation – Religion supports and nourishes scientific discovery.

The most important part of Haught's framework that will be considered in this discussion are the Conflict and Confirmation positions. According to Haught (1995 p. 10), the position of conflict is taken by scientists and skeptics (it could also be argued that religious literalists also adopt this position, just from the opposite side of the fence). Haught (1995 p. 11) says conflict arises due to the fundamentally separate ways religion and science obtain knowledge. In contrast, the confirmation position is not only reconciliatory but participatory too. As Haught (1995) puts it, "…the disinterested desire to know, out of which science grows and flourishes, finds its deepest confirmation in a religious interpretation of the universe (p. 22)." Haught is not alone in this position. John Polkinghorne (2011 p. 33), theoretical physicist and theologian, maintains a similar view that religion can inspire scientific enquiry. Though, it should be stated that, even without religion, people have been perfectly inspired to pursue scientific understanding..

The confirmation position is an agreeable position towards science. Unlike a young-earth creationist – who occupies the position of conflict – the arguments presented from confirmation will account for scientific discovery. But how exactly is this done? The next section will examine how the confirmation position grapples with the findings of modern science.

One of the ways theology has embraced modern science is in the very findings of science itself. Keogh (2015) says that scientific discoveries inevitably give rise to implicit statements and questions that can be answered through theology. Keogh (2015) argues that theology is positioned to answer questions of spirituality, existence, and meaning through a lexical function that science may lack. As an example of theologies lexical function, Haught (2011), arguing for the plausibility of immortality, says that science tells us that humans are a part of a bigger cosmological picture. And in this cosmic picture events can be thought of as being temporary, and as the universe unfolds events stack upon one another to preserve the past in the present (Haught, 2011). Haught (2011) goes on to propose that we should think of god, "as the repository of all the events that take place in cosmic history." Haught

(2011) claims that this argument cannot be proven by science, that means the corollary is also true: science cannot refute it. Haught takes an implicit assumption of science (events are temporary) and proposes a coherent interpretation of that through theology. This is precisely how Keogh envisions theology contributing to science.

Scientific Claims

Moving on from theology's proposed lexical functions, let us look at some direct scientific claims. Contemporary revelations in both physics and cosmology have shown that the universe relies on a collection of constants to exist in its current state. As Siegal (2015) explains: "as it turns out, it takes 26 dimensionless constants to describe the universe as simply and completely as possible…" In other words, if the force of gravity was too strong or the electromagnetic force too weak, the universe as we know it may have existed in an entirely alien way. A result of altering any one of the 26 constants could be that life as we know it would not exist. The discovery of these constants gave rise to the fine-tuning argument. However, we now know that the universe need not be as fine-tuned as was once thought (Adams, 2017). For example, take stars; they are far more robust than a fine-tuning purist would perhaps realize since the strength of the electric force could vary by a factor of 100 in either direction prior to any significant compromise of stellar operations (Adams, 2017).

Collins (2015 p. 213) maintains that a naturalistic explanation for these constants through further scientific theory is not enough, instead postulating that theism's response – that a god fine-tuned the universe – to be superior. Polkinghorne (2014) argues for this position too, saying that, "creation is not something [god] did fifteen billion years ago, but it is something that he is doing now (p. 73)." Justifying this stance, Polkinghorne (2014 p. 74) says that creation (or the Big Bang) can be separated from the creator, thus providing room for further ontological explanations regarding the origin of the universe (and attempting to avoid the logical question of 'who created the creator?'). This view is dependent on the interpretation of biblical scripture but essentially argues that after willing the universe into existence, god withdrew into himself; therefore, god does not come before creation but is apart of all creation (Polkinghorne, 2014 p. 74). Of course, further evidence for this position does overall appear to be lacking.

It is perfectly possible to be both a believer in god/s and science; though this does not make it intellectually or logically consistent. The confirmation position not only affirms science, but it also maintains legitimate inspiration for believing scientists. So, how do skeptics of the theology respond? We shall now continue our brief examination of the fine-tuning argument as an extension of Adams, 2017.

Is the universe really fine-tuned for life? Stenger (2011) argues that it is not. One solution for the seeming improbability of life can be found in current

cosmological models that suggest our universe exists as just one part in a sea of multiverses. In other words, our universe may not be so unique after all. Stenger (2011) acknowledges that this conclusion – which does away with the improbability of a life supporting universe – is disputed by theologians because it appears unscientific. How can we ever hope to observe a universe outside of our own? However, Stenger (2011) maintains, "a multiverse is more scientific and parsimonious than hypothesizing an observable creating spirit and a single universe." Another common objection to fine-tuning is that we lack an accepted definition of life (Freiderich, 2017). The notion that the universe is finely tuned for life presupposes knowledge of the conditions in which life can survive and the forms that life may take. Sober (2008 p. 77) presents a common objection to the fine-tuning argument known as the anthropic principle. The objection is as follows: since humans find themselves in a hospitable part of the universe, we are bound to think of those constants as being adapted for us; it is the observational selection effect in action (Sober, 2008 p. 77). Sober (2008) provides the following analogy to demonstrate what exactly the observational selection effect is: "Suppose you use a net to fish in a lake and observe that all the fish in the net are over 10 inches long. At first, this observation seems to favour the hypothesis that all the fish in the lake are more than 10 inches long over the hypothesis that only 50% of them are. But then you learn that the net has holes that are 10 inches across. This makes you realize that you were bound to obtain this observation, regardless of which hypothesis about the lake is true (p. 77)."

In summary, it is not all that strange to find ourselves in a hospitable part of the Universe on a relatively comfortable planet that orbits a relatively benign star.

Compatibility of Religion and Science?

A religious worldview is, on the surface, somewhat compatible with the scientific enterprise. This was demonstrated through both Haught and Polkinghorne. Confirmation theologians show a deep respect for scripture and science; often citing their religious belief as the inspiration for their scientific enquiry. This scientific reverence is shared by other noteworthy scientists like Francis Collins too, showing that religious belief has no real bearing on scientific enquiry. Indeed, none of Collins' great discoveries came about due to religion but thanks to the scientific method. In this sense, religion can motivate but not truly contribute. Indeed, even the natural philosophers of old such as Isaac Newton, Robert Boyle, Robert Hooke and Johannes Kepler, may have appealed to religious agents (that is, the supernatural), but always preferred and solidified their work with a scientific understanding. In the confirming position, there seems to be an admission regarding the validity of religion (perhaps to the point of cognitive dissonance). Scientific discovery offers humanity no guarantees or insights as to why something may exist. Its superiority as a method for understanding the world around us comes in part due to its falsifiability, and the lack of need for dogma or unwavering belief. And whilst the reasoning behind a phenomenon may not be known for years after a discovery is made, if at all – there is no reason to suggest or offer an alternative based on lexical presumptions. Science has made the need for guessing redundant. However, there may never be a satisfying answer for why the universe appears like it does, or why life exists, but scientific

enquiry does promise the chance for discovery. The need to know why may be spurred by the very nature in humans that needs to know how. But for now, science has cast a long shadow over the natural explanations and claims of religion. Theologies response appears to these authors as an extension of the god of the gaps argument. The god of the gaps argument (a type of logical fallacy known as the divine fallacy), states that mysterious phenomena that are currently unexplainable by science (or that science provides a currently incomplete understanding of) are explained by a divine being as the causative agent (Pennock, 2007). However, it falls short in imagination, as expressed by J. B. S. Haldane:

"Now, my own suspicion is that the universe is not only queerer than we suppose, but queerer than we can suppose. I have read and heard many attempts at a systematic account of it, from materialism and theosophy to the Christian system or that of Kant, and I have always felt that they were much too simple. I suspect that there are more things in heaven and earth that are dreamed of, or can be dreamed of, in any philosophy."

Religious Claims Investigated

Now that we have established how both religion and science are perceived, and the reasoning for the latter's superiority in understanding our world, we shall turn our attention to some particular religious claims.

Divine creation – or creationism – is a claim held dear to all three of the Abrahamic religions: Judaism, Christianity, and Islam (Ruse, 2014). In fact, one of the defining characteristics of a religion is that most have creation stories (Kurtz, 2015 p. 23). Creationism is the belief that the universe and everything in it was willed into existence by a creator-deity (Ruse, 2014). Creationism proposes an explanation for the existence of everything. But what does modern science have to say in contrast? Modern science sorts this investigation into various fields. The beginnings of life are investigated through chemistry and biology (a field known specifically as *abiogenesis*, which is distinct from *evolution*, which studies how this life changes over time), whilst planets and other stellar bodies are studied under the discipline of astronomy, with the fundamental workings of the universe questioned through the study of physics and its associated mathematics. Through these disciplines, people have discovered how life evolves in addition to developing the nebular hypothesis to account for the formation of stars and planets. Underpinning it all are the theories and laws of physics that let us pry into the very beginning of the universe itself. Although superficially there is no fundamental struggle between religion and science *per se*, religions have been forced to concede some of their faith claims. This is not unique to faith either, many scientific claims have been upended too, as Hawking and Jackson (2008) explain: "Newton's law of motions put an end to the idea of absolute position in space" (p. 18), and Einstein's theory of relativity abolishes the concept of absolute time. Science has a lot of pavers to lay and the garden is far from complete. Yet the scene before us has provided humanity with a far deeper understanding of the cosmos than any holy book could hope to provide. This is because science, as a system and method for understanding the world, is specifically designed to allow for

changes and updates via accumulation of bits of knowledge and Kuhnian revolutions. It is important to note that the crux of Thomas Kuhn's idea of scientific revolutions was that science does not just progress bit by bit, but by highly novel emergent ideas and theories, and that after exposure to scrutiny, cause an upheaval of the field altogether (Kuhn, 1996). Also importantly, this newly emergent information but be falsifiable for changes to be accepted and have any useful meaning (covered in more detail in Godfrey and Haggarty-Weir, 2017). Religions have no textual fundamentals that preach for this style of change, or even to facilitate it.

Holy books, unlike scientific texts, are full of miracles. The New Testament alone gives us water to wine, water-walking, and perhaps the most famous miracle of all, the resurrection of Christ. Lennox (2011) defines miracles as, "exceptions to recognized laws" (p. 167). As Hume (1793 p. 125) argued, not only do miracles violate natural laws, they also presuppose the existence of those laws. In that sense, miracles are best understood as suspensions of reality. Hume doubted the probability of miracles, arguing that it was rational to believe that in which was more probable to occur (Lacewing, 2014 p. 33). Could Jesus really walk on water or is it possible the story was exaggerated or simply made up? Hume's argument is not strictly scientific, but more an application of Occam's razor, a useful heuristic which states that when there are two competing theories the one that involves the least presumption is preferred. However, Hume's idea that probability could help uncover an objective truth is only made stronger by modern science. Take for example all that we know about water. Our intuition suggests that it would be difficult to walk on water. Indeed, as Pennisi (2014) explains: "humans are so big that the force of gravity overcomes the so-called surface tension of the water, making us sink." Indeed, it is through the understanding of buoyancy and surface tension that humans can build enormous ships or land spacecraft in the Pacific Ocean. Therefore, Hume would argue that, given what we know about water and floating things, postulating that a man could walk on it is so improbable that it likely never occurred.

This brings the discussion to another type of miracle, the idea of surviving death to lead an afterlife. The religious claim to an afterlife cannot easily be dismissed. However, the willingness of some religions to assign details to the afterlife is a claim to knowledge that no living being could possibly know – on the account that you would have to be dead in order to know it. The theological claim to an afterlife is often supported by near-death experiences (Mobbs & Watt, 2011 p. 447). However, Mobbs and Watt (2011 p. 449) argue that all parts of the near-death experience can be explained through neurophysiological means. Even our experience of death – or what we perceive as death – takes place inside the brain. Russell (2013 p. 4) argued that our mental life – or soul – is entirely reliant on the working functions of the brain. Therefore, when our body stops functioning it is probable to suppose that the soul ceases with it (Russell, 2013 p. 4). Given the evidence, the idea of an afterlife does not seem to be supported by anything except the claim itself. Let us for a moment apply the simple heuristic of Occam's razor; why add more to explain something than is fundamentally required for understanding of said phenomenon? Making the case to demystify the afterlife, Mobbs and Watt (2011)

conclude that only when science begins to probe these questions can we, "move beyond theological dialogue and into the lawful realm of empirical neurobiology" (p. 449).

The faith claims of religion may still hold some intrinsic merit. Their value should not be entirely dismissed because they are not supported by empirical evidence. Anticipating her final departure, the widow may find comfort in the belief that her spouse awaits on the other side. Of course, it could be argued that the opposite is also comforting: no long-dead spouse awaits your arrival. Which is better? The futile anticipation of being reunited with a loved one, or the final realisation that you were lucky to share a life with those whom you loved?. Of course, both thoughts are compatible, but it is only the faithless who forego first. Further, one might argue there is intrinsic intellectual dishonesty in holding a belief solely for comfort. Faith claims provide comfort for many, but it is the scientific mode of inquiry that truly demystifies the world. It should be therefore argued that more people should undergo a paradigm shift and find an equivalent level of comfort in the scientific facts. Yes, our loved ones may never be seen again, but they return to their proverbial cosmic ether, and you will too one day; there is a level of beauty in this idea and it does not require belief in the indemonstrable.

Conclusion

It seems less plausible that there is new knowledge that can be obtained through religion that cannot be obtained through science, and the opposite of that is not likely to be true. Indeed, modern science has certainly cast a long and permanent shadow over the faith claims of religion. But this new figure is not to be feared because, as Russell (2013) puts it, "even if the open windows of science at first make us shiver after the cosy indoor warmth of traditional humanizing myths, in the end the fresh air brings vigor, and the great spaces have splendor of their own" (p. 7). And finally, the authors posit a concluding question to those of the religious persuasion that now accept the role and superiority of scientific reasoning for understanding the world around us, but that still maintain a role of religious reasoning in other areas of life: couldn't you not simply replace the religious reasoning with another philosophical reasoning (i.e. ethics) so as not to require the additional complexity (and superfluousness) of invoking the supernatural? This is something for readers to be left to contemplate honestly on.

Acknowledgements

The authors would like to give thanks to Joshua Cox from the University of Edinburgh for his insightful comments during the drafting of this paper.

References

Adams, F. (2017). The Not-So-Fine Tuning of the Universe. *Cosmos*, Nautilus.

Ayala, F. J. (2009). Darwin and the scientific method. *Proceedings of the National Academy of Sciences, 106*(Supplement 1), 10033-10039. doi:10.1073/pnas.0901404106.

Collins, R. (2012). The Fine-Tuning of the Cosmos *The Blackwell Companion to Science and Christianity* (pp. 207-219): John Wiley & Sons, Ltd.

Feynman, R. P., Robbins, J., & Dyson, F. (1999). The pleasure of finding things out: The best short works of Richard P. *Feynman. Cambridge, MA: Helix/Perseus*.

Friederich, S. (2017). Fine-Tuning. Fall 2017. Retrieved from https://plato.stanford.edu/archives/fall2017/entries/fine-tuning/

Godfrey, W. & Haggarty-Weir, C. N., 2017. Popper's Formulation of Scientific Knowledge: A shift from an inductivist account. *The Philosopher*, June.

Haught, J. F. (1995). *Science and religion: From conflict to conversation*: Paulist Press.

Haught, J. F. (2011). Science, self, and immortality.(Report). *Annals of the New York Academy of Sciences, 1234*, 70.

Hawking, S. W., & Jackson, M. (2008). A brief history of time: Bantam New York, NY.

Hume, D. (1793). *An inquiry concerning human understanding* (Vol. 3).

Keogh, G. (2015). The Value of Theology in a Scientific Age. *New Blackfriars, 96*(1064), 405-418.

Kuhn, T.S. (1996). The Structure of Scientific Revolutions. 3rd ed. Chicago, IL: University of Chicago Press.

Kurtz, L. R. (2015). *Gods in the global village: the world's religions in sociological perspective*: Sage Publications.

Lacewing, M. (2014). *Philosophy for A2: Unit 4: Philosophical Problems*: Routledge.

Lennox, J. C. (2011). *Gunning for God: Why the new atheists are missing the target*: Lion Books.

Mobbs, D., & Watt, C. (2011). There is nothing paranormal about near-death experiences: how neuroscience can explain seeing bright lights, meeting the dead, or being convinced you are one of them. *Trends in cognitive sciences, 15*(10), 447-449.

Pennisi, E. (2014). The Secret to Walking on Water Retrieved from http://www.sciencemag.org/news/2014/03/secret-walking-water

Pennock, R.T. (2007). God of the gaps: the argument from ignorance and the limits of methodological naturalism. Retrieved from https://s3.amazonaws.com/academia.edu.documents/38038401/PennockGod_of_Gaps.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1528646558&Signature=jxwqqE4PUB%2ByTLiGCq1w3I0u%2Fpo%3D&response-content-disposition=inline%3B%20filename%3DGod_of_the_gaps_The_argument_from_ignora.pdf

Polkinghorne, J. (2011). *Science and Religion in Quest of Truth*. New Haven, UNITED STATES: Yale University Press.

Polkinghorne, J. C. (2014). *The Faith of a Physicist Reflections of a Bottom-Up Thinker*. Princeton: Princeton : Princeton University Press.

Ruse, M. (2014). Creationism. Summer 2014. Retrieved from https://plato.stanford.edu/entries/creationism/

Russell, B. (2013). *What I believe*: Routledge.

Siegal, E. (2015). It Takes 26 Fundamental Constants To Give Us Our Universe, But They Still Don't Give Everything. Retrieved from

https://www.forbes.com/sites/ethansiegel/2015/08/22/it-takes-26-fundamental-constants-to-give-us-our-universe-but-they-still-dont-give-everything/#518580974b86

Sober, E. (2009). Absence of evidence and evidence of absence: Evidential transitivity in connection with fossils, fishing, fine-tuning, and firing squads. *Philosophical Studies, 143*(1), 63-90.

Stenger, V. J. (2011). *The fallacy of fine-tuning: Why the universe is not designed for us*: Prometheus Books.

Statistical Hypothesis Testing in the Context of Hume's Critique of Induction

Leo Lobksi

Perhaps the best-known formulation of the problem of induction [5] can be found in David Hume's *Treatise of human nature* [1, Book I, Part III, Section VI]. There, Hume draws our attention to arguments of the following form: we first repeatedly observe some contingent property for some objects (without observing all the objects), and if we discover that all the observed objects have the property, we conclude that *all* the objects (including those which were not observed) have this property. A weaker form of this argument would be to conclude that it is *more likely* that the objects in question have this property. As an example, suppose that we are examining books on a bookshelf one at a time, having gone through half of the books, we discover that all covers thus far have been green; reasoning inductively, we conclude that all the books on the shelf are green. As another example, we can observe that the first five even numbers greater than two can be written as a sum of two prime numbers: $4 = 2 + 2$, $6 = 2 + 3$, $8 = 5 + 3$, $10 = 5 + 5$, $12 = 7 + 5$; we are thus tempted to postulate that the same is true for any even number greater than two. This type of argument is, of course, fallacious; there is simply no reason why observing something about a part of a larger collection would imply that the same is true about the entire collection. Or as Hume puts it [1, I.III.VI, p. 88],

> *"From the mere repetition of any past impression, even to infinity, there never will arise any new original idea, such as that of a necessary connexion; and the number of impressions has in this case no more effect than if we confin'd ourselves to one only."*

My main objective here is to argue that, despite this is seemingly impractical, we should take Hume's criticism seriously. Furthermore, I propose that the framework of hypothesis testing in statistics accomplishes this; I suggest to view this as a factor contributing to the success of statistics as the foundation of contemporary empirical sciences. I will, however, unlike Hume, argue that under certain strict conditions an inductive inference can be correct, provided that we relax 'correct' to mean 'extremely unlikely to be false'. Although Hume connected the problem of induction to that of causality, and there are no doubt interesting connections between the two, the focus here will be on the form of the inductive argument.

What makes Hume's account of induction highly problematic is its evident conflict with practice. Indeed, inductive reasoning is ubiquitous both in everyday life and in the natural sciences; if the bridge has been strong enough to cross thus far, it will be safe to cross in the future; if no stone has been seen floating on water, then no stone at all floats; if things have always fallen down, perhaps they will continue to fall down. Although in our introductory examples the reasoning is clearly invalid, there are many cases in which it seems that inductive inference is justified. Say we are boiling pasta, and we checked that two pieces of pasta are well cooked, from this we conclude that all the pasta in the pot is ready to eat. In this case, there seems to be a

good underlying reason to assume that if two pieces of pasta are well cooked, then so are the other ones. Such reasons could, for instance, include the assumptions that temperature everywhere in the pot is the same and all pieces of pasta take the same time to cook. In this example, it is important to note that although intuitively the reasoning seems correct, it is formally invalid; it is conceivable that, for example, the two pieces of pasta we tried just happened to be thinner than all the other ones, and moreover happened to be closer to the hob, and were hence cooked faster than the other pieces. Since this possibility exists, *logically* we are not allowed to conclude that the pasta is cooked until we have eaten all of it.

Inferences based on reasoning similar to that in the cooking example are so mundane that we make them without even paying attention. More than that, they are routine in the scientific practice. Take the example of Newtonian gravity; there it is first observed that massive objects exert an attractive force on each other (note that this can only be measured for objects with a large mass which are not too far apart from each other), which is then generalised to any objects with mass at any distance from each other. This is a courageous step which does not formally follow from any observations. Importantly, the inductive inferences, both in everyday life and in the sciences, tend to be blatantly successful. The pasta will indeed be well cooked, the bridge will be safe to cross, objects with mass will attract each other etc. The conclusion is, therefore, that the problem we face is the disparity between the success of inductive reasoning and the lack of justification for it. Hume recognised this disparity, but he insisted that there is no justification for induction in principle. Instead, he explained inductive inferences as some kind of practical necessity, a habit human beings engage with because we are so structured, with no other reason or justification for it (see e.g. [1, I.III.XII]). This is, however, highly unsatisfactory in the light of success of inductive inference in the natural sciences, which supposedly reach beyond mere human habits [6].

One possible position on induction is to insist, like Hume, that inductive reasoning is always wrong, and that there is always some other reason why the inference is valid when induction happens to produce the correct outcome. This certainly captures the essence of the mathematical example given; while inductive inference can be used to get a hypothesis or a conjecture, a proof is required to show the conjecture true or false. We can perhaps get away with the same attitude when considering the bridge example too; it is not sufficient that the bridge was safe to cross before, it still has to be in a good condition to remain safe. What this position of always requiring a non-inductive inference does not capture are situations with random sampling. This brings us back to the example with potentially raw pasta. In this case, it seems, induction is the only type of inference we can possibly use, unless we eat all of the pasta, at which point it becomes irrelevant to know whether it was cooked or not. It appears that the inference is truly inductive, as one can always come up with some anomaly falsifying any deductive argument deriving aldenteness of the pasta from that of two pieces. Yet it also appears that this inference ought to be, if not valid, then at least probabilistically valid. With these observations in mind, let us analyse the logic of this argument. The reasoning here goes somewhat as follows: the scenario in which only some part of the pasta is cooked is already rather unlikely; then, even if we are so unfortunate that the unlikely scenario happened and some

pasta is still raw, when we randomly take the two pieces of pasta from the pot, with some probability, at least one of them will be uncooked, which would be enough to conclude that not all pasta is cooked. Overall, the probability of some pasta being raw *and* us taking two well cooked pieces of pasta out of the pot is very small. Hence, in the case that we do indeed take out two well cooked pieces, we can safely reject this scenario. Here 'safely' means 'with a low chance of being wrong'. The certainty with which we can reject the scenario can of course be increased by taking out more pieces of pasta. Crucially, this reasoning does not rely on a fallacious inference, at least under the understanding that when we say 'the pasta is cooked' we mean that it is extremely unlikely that some of it is still raw. This is not unreasonable usage of language, as with a sufficiently large pasta sample the frequency with which a false inference would occur is practically zero; the number of times one has to cook pasta in order to observe this scenario perhaps exceeds the total number of pasta pots ever cooked [7].

The lesson learned from the example with cooking pasta is that in order to avoid invalid inferences, and if we insist on taking Hume's criticism of induction seriously (which we do), we are forced to enter the world of probabilities. This brings us to hypothesis testing in statistics. We already described the basic idea of a hypothesis test via testing the hypothesis 'some pasta is raw'. Without going into too much detail, we will next describe a general hypothesis test.

Statistical hypothesis testing concerns itself with a setup when there is some set whose property we are interested in (e.g. all the books on the shelf, all the pasta in the pot, all people in the world), this is called the *population*. It is assumed that we do not have access to all of the population; this could be due to the size of the population (people in the world), because we do not yet have (or choose to ignore) the complete information about the population (bookshelf), or because testing the entire population would be impractical (people, pasta in the pot); the reason for not having the access is, however, irrelevant. What we do have access to is some subset of the population, called a *sample*. Given a sample, we can test a hypothesis about the entire population as follows. We compute what is the probability of getting the sample as observed (or a more extreme result) assuming that the hypothesis is true, this probability is called the *p-value*. In short:

$$\text{p-value} = P(\text{S or a more extreme result, assuming that H holds}),$$

writing P for probability, S for sample and H for hypothesis [8].

Prior to computing the p-value, we set a significance level α, which is taken to be some 'small' positive real number. If the p-value is less than α, we conclude that it is so unlikely to get the given sample if the hypothesis is true, that we can reject the hypothesis in favour of another one yet to be tested. It is, of course, possible to reject the true hypothesis just by chance; in order to gain further certainty that a hypothesis can indeed be rejected, one can lower the significance level or increase the sample size. The significance level is the rate at which we would reject the hypothesis H while H is in fact true if we repeated the test many times. This explains why we want α to be 'small'. It is important to note that the meaning of 'small' depends on the context

and on the desired certainty; while $\alpha = 0.05$ or $\alpha = 0.01$ are often used, for example, for the CERN experiment discovering the Higgs boson, the significance was as low as one in three million [4].

It is crucial for the validity of the argument in a hypothesis test that the probability is computed *assuming* the hypothesis is true. This is precisely how inductive argument is avoided; at no point we infer that the hypothesis is true by simply observing a sample. Instead, we ask if the hypothesis were true, with what probability would we see such a sample. One implication of this is that we can never confirm the hypothesis (except in some cases when the hypothesis is something quite simple, as in, H = 'Not all books are green', and we find at least one book which is not green). The terminology used is that we either 'reject' or 'fail to reject' the hypothesis at the significance level α. Why this clumsy terminology is used is explained by the cooking pasta example again. If we chose the significance level low enough, we would fail to reject the hypothesis H_1 = 'At least one of the pieces of pasta in the pot is still raw', hence at this ridiculously low significance level we cannot distinguish between H_1 and H_0 = 'All of the pasta is cooked', whose p-value is 1 if we have indeed two well-cooked pieces of pasta in our sample. Arguably though, the significance level for determining whether the pasta was cooked or not does not need to be this low. Because of this ambiguity, many authors insist that truth of a hypothesis can only be tested against an alternative hypothesis (see [3, 2.2.] or [2, ch. 6&7]).

It is also worth emphasizing that the setup of hypothesis testing captures our intuition that there must be some underlying reasons justifying the inference from a sample to the entire population. Namely, in order to compute the probability of getting a result at least as extreme as the sample given some hypothesis H, we need to know in what way the truth of H affects the distribution of outcomes.

To make the discussion more concrete, let us return to the bookshelf example. Let us take the total number of books (population size) to be 100, and suppose that we have gone through 50 books (sample size). Consider the simple hypothesis H_0 = 'All books are green'. If the sample contains at least one book which is not green, there is nothing to test, hence suppose all 50 books in the sample are green. Then, the probability of this outcome assuming H_0 is of course 1, and so we fail to reject H_0 . Note that, as in the remark above, it does not follow that H_0 is true. Compare this to another hypothesis, H_1 = 'Exactly one book is not green', assuming that any ordering of the books on the shelf is equally likely, the p-value, i.e. the probability that all 50 books in the sample are green given H_1 , i.e. the probability that the non-green book is in the other half of the shelf, is 0.5, and so we fail to reject H_1 as well (at any reasonable significance level). All this is saying is that the observed sample is consistent with both assumptions. If we, however, consider H_2 = 'Exactly 10 books are green', then (again, assuming that any ordering is equally likely), then the p-value is approximately 0.00059342, and so H_2 will be rejected at significance level α = 0.001. Here we are very strongly using the assumption that any ordering of books is equally likely (the underlying reason justifying the inference); if we had any reason to suppose this is not the case (e.g owner of the shelf likes to arrange their books according to colour), the p-values and hence the conclusions would change.

The hypothesis testing was presented here in the simplest possible form, avoiding technicalities. For a much broader discussion the reader is referred to Hacking's *Logic of statistical inference* [2]. The emphasis here has been on general logic underlying all hypothesis testing rather than on any particular details, and in particular, on how this logic relates to the problem of inductive inference. I by no means claim that hypothesis testing solves the induction problem; the inductive inference, of course, remains invalid as demonstrated by numerous examples above. Furthermore, one is forced to accept Hume's position that there is no way induction can ever be valid or justified, unless we interpret validity of the inference probabilistically, as we have done here. Even when we relaxed our criterion for validity, we noticed we had to make a lot of assumptions in order to justify induction. Hence in most cases Hume's criticism remains relevant. The power of hypothesis testing lies in *accepting* this criticism; it replaces induction with a hypothesis postulated *prior to* rather that after the observation. Note that it does not matter at all how we arrive at the hypothesis, it could be a guess, induction, intuition, divine providence or dark magic, it makes no difference to its status as a hypothesis whatsoever. Despite this, hypothesis testing manages to fill the gap between the plausibility of certain inductive arguments and the lack of formal justification for them. Not only that, but it also explains the success of statistics as a foundation of modern empirical sciences. The success lies precisely in taking the unjustifiability of induction seriously, in replacing the inductive inference with something logically sound, while constantly keeping in mind and quantifying the possibility of error.

References

1 Hume, David. A treatise of human nature. Oxford at the Clarendon Press. 1888. ISBN 0198241305.

2 Hacking, Ian. Logic of statistical inference. Cambridge University Press. 1965. ISBN 9781107144958.

3 Mayo, Deborah G. Spanos, Aris. Severe Testing as a Basic Concept in a Neyman-Pearson Philosophy of Induction. The British Journal for the Philosophy of Science, vol. 57, no. 2, 2006, pp. 323357. JSTOR. jstor.org/stable/3873470.

4 Observation of a New Particle with a Mass of 125 GeV. CMS experiment, CERN. cms.web.cern.ch/news/observation-new-particle-mass-125-gev.

5 Vickers, John. The Problem of Induction. The Stanford Encyclopedia of Philosophy (Spring 2016 Edition). Edward N. Zalta (ed.). plato.stanford.edu/archives/spr2016/entries/induction-problem.

Footnotes

[6] It is debatable whether an outcome is confined to the domain of the method. Although the scientific method can be a human habit (e.g. inductive reasoning), the resulting scientific theory may well contain elements not contingent on human

perception. This does not avoid the plausibility problem though, we could equally formulate the question as why these particular human habits rather than some others lead to successful theories.

[7] This is a wild guess, although it must be true for a sufficiently large sample. And in any case, the frequency will be extremely small.

[8] Here ‘or a more extreme result’ is included for technical reasons, as otherwise any single outcome would be extremely unlikely. If in our pasta example the sample S consisted of one cooked and one uncooked piece of pasta, and the hypothesis was H = ‘ one third of the pasta is uncooked’, a more extreme result would be two pieces of uncooked pasta. This plays no significant role in understanding the basic logic of a hypothesis test though.

The Prudential Value of Self-killing: *What conditions could ever make the cessation of the self appealing?*

Kenneth Novis

Discussion of suicide is understandably controversial. Whilst Camus famously claimed that there is "one truly serious philosophical problem, and that is suicide",[12] David Benatar in his discussion of this topic prefaces that "One cannot write in support of (some instances of suicide) without considering the possibility that some desperate person might read one's words and ... [end] his or her life."[13] With this sentiment in mind, this topic is rightly controversial – anyone failing to acknowledge this fact does so at great risk to their readership. I do not mean to endorse suicide; what this essay aims to achieve is an accurate portrayal of arguments regarding suicide, so that readers of it may find their own ways to disagree and triumph over pessimism as works best for them. In this sense, writings such as this are never pleasant ones (instead dealing with the very depths of human despair) but are nonetheless among the most important: to lay the arguments bare, make them accessible to public scrutiny and, hopefully, dismissal. With this in mind, I ask you to proceed highly critically, and to heed the cautionary message that I do not convey this information lightly.

In his most recent publication *The Human Predicament*, the philosopher David Benatar flits around a question to which this essay will be devoted: what conditions could ever arise to tempt a person to commit self-killing? In this article, I mean to neither repudiate nor merit the act of self-killing; instead, my discussion will be effectively eschewed from the normative claims surrounding the topic. A discussion on the grounds of prudential appraisal will be engaged in – a tradition seeking not to apprehend the morality of a situation, but rather to consider what might be done to increase someone's well-being, or make their lives go better.[14] Thus, this shall revolve around the soundness of motives to commit self-killing, and not the morality. Whilst Benatar's response offers a compelling resolution to the issue of self-killing, he does not do so with the due care for prior existential work on the topic equivocating his claims, nor is his work of advocacy thorough in downplaying the rationality of a more hopeful perspective than that which he espouses, such as that of the progressive optimist (who believes that there are sufficient reasons to be hopeful of the future) and the pragmatic optimist (who believes that, even if pessimism is empirically sound, it exacerbates the harm of pessimism to accept its thesis). Here,

---

12 Camus, A., O'Brien, J. (2005) An Absurd Reasoning, *The Myth Of Sisyphus (Penguin Great Ideas)*. Penguin Group: London. Pg. 1.

13 Benatar, D. (2017) Suicide, *The Human Predicament*. Oxford University Press: New York. Pg. 165.

14 Weijers, D. (Last viewed 25/11/2017) Value and Prudential Hedonism, *Internet Encyclopaedia of Philosophy*, forthcoming URL = <https://www.iep.utm.edu/hedonism/#SH1b>

pessimism is taken to be the rational expectation of the worst over the best. I have structured this essay by establishing a brief groundwork for acceptance of self-killing *sometimes*. I then commit to undermining views of pessimism as irrational, or as an empirically unsuitable alternative to progressive optimism. Following this, I regard existential questions of meaning, and the extent to which they may affect well-being, finding responses to these questions by pragmatic optimists unsatisfactory. Finally, I conclude in summary: noting that pragmatic optimism fails to resolve the complaints lain before it; that there is a strong rational precedent to pessimism; and that proposed resolutions to existential concerns in pessimism are either ineffective or exclusive in their efficacy.

First on our call of order is the need to establish grounds on which we might accept self-killing *sometimes*. If this was not firstly established, then there could be no sufficient reason for which one would feel compelled to take their own life; self-killing being at the very least reasonable is then of primary value to this discussion. In Benatar's view, self-killing is rendered a reasonable alternative once one is destined for a fate "worse than and only avoidable by death"[15] and so, contrary to previous philosophers such as Hobbes who have argued at length that no act failing to preserve the self can ever be justified,[16] on rare occasion the act in greatest service of the agent's own interests may be the cessation of the self.

Progressive optimists would generally contend that because a person committed to self-killing will be doing so from a compromised epistemic stance, eroded perhaps by depression or anxiety, they cannot validly condemn their whole future with the belief that things will never become better. It may well be a delusion of the depressed that life, when despondent, is unlikely to recover. If true, then the logical stance would only be that self-killing is warranted in the interests of maximising prudential value when the agent is sound of mind and so self-killing may still be positive for well-being sometimes. For even the sound of mind, it may be in their interests to evade a "fate worse than and only avoidable by death". Two additional responses to optimism will be further considered: (1) The conditions of human life lend themselves to the rational expectation of all manner of future hardships. If true, then whilst we oftentimes rightly dismiss the depressed for their aberrant world view, it is not always irrational to hold such an appraisal as that 'conditions will never improve'. (2) That instances exist where the events leading to the onset of such a pessimistic worldview cannot, by their nature, become better, regardless of the efforts of the agent. In the first of these conditions, the agent is rational to hold a pessimistic worldview, and in the second, the agent is rendered incapable of electing anything other than such a worldview by nature of its etiology. Both of these responses counter the view that 'all acts of self-killing by the unsound of mind are invalid, were it the intention of the agent to maximise their well-being in doing so.'

The first of these two views is prone to one's beliefs regarding life: does one believe it to be more abundant in pleasure or hardship? Which of these ought to be

---

[15] Benatar, D. (2017) Suicide, *The Human Predicament*. Oxford University Press: New York. Pg. 163.

[16] Hobbes, T., Gaskin, J. (2008) Of Man, *Leviathan (Oxford World Classics)*. Oxford University Press: New York. Pp. 88-89; 93-95.

given precedent when making decisions within life? It's easy and effective to be an optimist in this regard: Joshua Miller in *'Utopic pessimism: The messianic underpinnings of the antinatalist polemic'* confesses that for reasons of happiness, one might wish to willingly neglect the pessimistic stance;[17] that doing otherwise consigns one to a realm of utter impotence, within which they are passive agents in the inevitable self-destruction of all humanity. Given the bleak and hopeless outlook of pessimism, it may seem pragmatic to reject it. However compelling a view this one of hope is, one must also consider the mindset of such persons as are likely to conceive of self-killing as a valid alternative to continuation of existence – it is their motivations and underlying compulsions which we sought out initially to analyse regardless.

To psychologically healthy people, the doxastic venture necessary to move from the burdensome empirical weight of pessimism to the wholehearted acceptance of optimism may be a simple stride away from a view better abjured than endorsed for pragmatic reasons, to one potentially more practical in its implications, but empirically questionable. Although we shall later interrogate the veridicality of pessimism, this move (provided that pessimism withstands testing against reality) is itself a rational response. However, depressed persons may not be so well equipped to shrug off this weight. Whilst, for the already healthy it is easy to deafen oneself to the empirical claims of pessimism, for those already handicapped by their affectations, they may be harder to deflect (in part also for reasons in the second view, discussed later on). For this reason, a hopeful rhetoric recited to the deeply depressed, such as to say "cheer up, things aren't that bad" may do more evil than good. Not only are they already rendered incapable of picking themselves up in the same way, or just 'snapping out of it', their condition may be worsened by insistence that they should be happy, and thus that something is profoundly at fault in them.

Whilst I have taken it as granted that pessimism has a strong empirical grounding, we ought to consider that optimism may also, and thus the belief that life will always remain of poor quality would be irrational by all accounts. The pessimistic claim is that life gives more cause to consider it negatively than positively: the greatest probability is that any human being born will begin life in utter squalor, with limited to no satisfaction of their basic needs (food, water, shelter, etc.); further, no life is without hardship, and one of the few guarantees within existence is the gradual decline in conditions of life from childhood to adulthood to old age – the diseases and afflictions contracted over the course of a person's life can be cripplingly painful, or leave one in a disabled state, and lastly, whatever good or pleasure that does exist is exceptionally fleeting, as opposed to the suffering inflicted upon people, which is almost universally longer lasting in effect. There exists no 'optimistic alternative' to a broken arm, or a long-lasting and debilitating disease, short of simply not having either affliction (which is no cause for pleasure, it is only an experience of normalcy).[18]

---

[17] Miller, J. et al. (2015) *Utopic pessimism: The messianic underpinnings of the antinatalist polemic.* ProQuest Dissertations and Theses. Pp. 63-64

[18] Benatar, D. (2017) Quality, *The Human Predicament*. Oxford University Press: New York. Pg. 77

Contrary to this, the philosopher John Dewey claims that there do exist sufficient reasons to be hopeful for the future; that there are sufficient goods in life; that these goods justify our expectation of more future goods.[19] This view he calls 'Meliorism', and for the first of its three claims there seems to be overbearing empirical evidence. The quality of life now is vastly better than last century, with decreases in infant mortality rates and basic sanitation having progressed monumentally. However, despite the historical rapidity of such developments, they are unlikely to make noticeable advancements over the course of a single life, and what progress they do make is unlikely to be satisfactory to a person already considering self-killing. It's difficult to see how these advancements could be guaranteed to remedy every affliction which the agent might experience (as any sufficient malady may validate prudential self-killing) or downplay every hardship they might confront when they are also likely to offer no new solutions whatsoever. The optimist's gamble is to assume that science will progress in a way to reduce specific suffering and the possibility of return to normalcy (which does not in itself elicit pleasure) when the gamble of the pessimist is to assume that science will make no such specific progress, and thus self-killing serves to remove a suffering which may never be resolved. More simply, there exist 4 possible outcomes and associated prudential value, represented in figure 1.0.

	Self-killing	***No self-killing***
Conditions improve	Bad *(Pleasure is missed out on, but so is all future suffering)*	Not bad *(The agent returns to a normal state)*
Conditions do not improve	Good *(There is a reduction in their experienced suffering)*	Bad *(Their suffering does not subside)*

Figure 1.0 shows that whilst not self-killing can lead to a reduction in suffering or its persistence, self-killing can only ever reduce net suffering, albeit at the potential loss of some pleasure.

Figure 1.0 is roughly abridged from Benatar's *Better never to have been: the harm of coming into existence*,[20] but against it there is a final defense the optimist may rouse to maintain the irrationality of pessimism: it could be said that it is more than *'Not bad'* when someone lives to experience their quality of life recovering from severe suffering; that this is a marked increase in quality of life. Because prudential value considers what things make a life go well or better, recovering from suffering is of great worth to it. If someone was able to eliminate suffering from their lives, it ought to be characterised as more than a 'return to normalcy', but instead to be beneficial to the person. However, it is beyond the grasp of modern medicine to relieve all afflictions, physical and existential – under modern conditions, and with

[19] Shade, P., & Lachs, John. (1997). *Habits of Hope: A Pragmatic Theory of the Life of Hope,* ProQuest Dissertations and Theses. Pg. 22.

[20] Benatar, D. (2008) *Better Never to Have Been: The Harm of Coming into Existence.* Oxford University Press: New York. Pg. 39

the current trajectory of science, it is far less likely that a person living presently could live a life free from suffering, and given that suffering exists in much greater magnitude than pleasure, one cannot cling to vague hope to annihilate all suffering, only to potentially minimise it to an unknown extent. Given this evaluation of the current conditions of science and suffering, it is more rational to remain pessimistic than hopeful or optimistic.

Whilst the first concern dealt with the rationality of pessimism, the second ought to show that, once conceded to, a certain sort of existential pessimism is immutable, and that this existential pessimism is much more gravitas in nature than Benatar credits it as being.[21] For Schopenhauer, meaninglessness resides in the ironically endless pursuit of ends, and in this pursuit (deemed by later existentialists 'the absurd') people may come to lament the meaninglessness of their labour.[22] This view is simplified by Benatar, describing life as something which seems to be comprised of many simple activities with little greater end than to generate further meaningless tasks.[23] This being the crux of meaninglessness, it would seem that the realisation of it leads to confrontation of an absurd existence: one where all acts have fed into this perpetual cycle, and can rationally be expected to continue as such because all human acts do so.

Camus in '*The Myth of Sisyphus'* discusses this loss of meaning, confessing that most anything which disrupts our views of the world and reveals the underlying absurdity can lead to experience of ennui, or disenchantment with life's functions.[24] However, an important feature of existentialism is that once uncovered, the illusions of objective meaning and worth are beyond recovery. Philosophers have responded to this in a variety of ways, from Nietzsche positing that one ought to forge their own subjective meaning,[25] to Camus' belief that one ought to remain positive in spite of it.[26] However, both of these stances seem to be negligent of the very nature of this awakening: that the realisation that no non-arbitrary meaning exists leads one to despair and crave the return to objective value, when both philosophers are proponents of a reconstruction of meaning in overtly subjective terms. It seems unlikely that, when angst arises from realisation of the subjectivity of meaning, it could be remedied with only subjective meaning also.

---

21 Benatar, D. (2017) Suicide, *The Human Predicament*. Oxford University Press: New York. Pp. 190-194

22 Schopenhauer, A., Norman, J. (2010) *The World as Will and Representation, Volume 1*. Cambridge University Press: Cambridge. Pg. 389

23 Benatar, D. (2017) Meaning, *The Human Predicament*. Oxford University Press: New York. Pp. 14-15

24 Camus, A., O'Brien, J. (2005) An Absurd Reasoning, *The Myth Of Sisyphus (Penguin Great Ideas)*. Penguin Group: London. Pp. 2-3.

25 Nietzsche, F., Hollingdale, R. (1974) The Three Metamorphoses, *Thus Spoke Zarathustra.* Penguin Group: London. Pp. 54-55; Nietzsche, F., Kaufmann, W. (1991) The Madman, *The Gay Science.* Random House Inc.: USA. Pp. 181-182

26 Camus, A., O'Brien, J. (2005) The Myth of Sisyphus, *The Myth Of Sisyphus (Penguin Great Ideas)*. Penguin Group: London. Pp. 115-119.

It is possible that these solutions are offered not because they successfully resolve the issue of cosmic meaninglessness, but rather because they are all that one can do to continue the illusion of meaning; that, to use Camus' phrase, one "should imagine Sisyphus happy",[27] or otherwise remain deeply unhappy themselves. However, Benatar deals with meaninglessness with less sincerity than he perhaps ought to in believing that meaning is prudentially unimportant.[28] In his view, whilst the absence of satisfactory meaning in a person's life can compel them to end their lives, it is never this factor alone – in his defence, it is difficult to imagine a life otherwise going exceptionally well, save for the agent's experience of ennui, where they do commit an act of self-killing, as he rightly notes. It does seem to be the case that meaninglessness is only ever an auxiliary cause of self-killing; people who kill themselves have many things wrong with their lives, the least of which is generally meaninglessness.

This view downplays the sincerity of existential claims, however: Cioran notes that "life is only possible by the deficiencies of our imagination and our memory".[29] In essence, that had we all an accurate appraisal of life as it exists, and an accurate perception of all things which have happened to us, nobody would want to endure in this existence. It is, then, only by certain deficiencies that we fail in possessing the logic to see the despondency of our lives, become wholly disillusioned of them, and therein lose all joy in living them. This final point is the critical one which I hoped to raise: whilst meaninglessness may not be as horrific in itself as torture or terminal disease, it is likewise difficult to imagine someone taking joy from something without first deriving something *meaningful* from it. One might imagine that all manner of activities, once deprived of meaning per this awakening no longer derive the same joy as before, them being merely constituents of the hamster-wheel of human activity. Nothing being meaningful, or providing meaningful pleasure, might be likened to deriving no pleasure from things whatsoever, and so ennui has the much greater potential to reduce well-being than Benatar believes – instead, it denigrates all good things in life, once it is understood that they are all arbitrary.

The condemnation of pragmatic optimism is also of note here: Camus' response may function, were one able to blind themselves to this cosmic meaninglessness, but for those incapable of doing so, who unwillingly see the nihil wherever they look and without hope of recovery from it, prospects for recovering the joy in life are diminished. The classical solution offered against ennui is to forge meaning for oneself, or otherwise become existentially blind, and eschew oneself of all concerns regarding meaning – to become Camus' 'absurd hero',[30] but this is an optimism only suited for some, and many still exist for whom the epiphany of the

---

[27] Camus, A., O'Brien, J. (2005) The Myth of Sisyphus, *The Myth Of Sisyphus (Penguin Great Ideas)*. Penguin Group: London. Pg. 119.

[28] Benatar, D. (2017) Suicide, *The Human Predicament*. Oxford University Press: New York. Pp. 190-194

[29] Cioran, E., Howard R. (2010) The Key to Our Endurance, *A Short History of Decay (Central European Classics)*. Penguin Group: London

[30] Camus, A., O'Brien, J. (2005) The Myth of Sisyphus, *The Myth Of Sisyphus (Penguin Great Ideas)*. Penguin Group: London. Pp. 115-119.

nihil isn't easy to deactivate. It is for this reason which I do not believe that Benatar justly downplays existential concerns of meaning, nor that these existential concerns, once alight are easily doused with pragmatic optimism: it may be an option, albeit one only available to those with already the strength of character to seize themselves from the depths of knowing despair and migrate to the heights of blind joy; this subset of people will rarely if ever include those already considering self-killing, and because this change of perspective does not resolve the issue, only involves becoming numb to it, the insistent pessimist will still rightly note that meaninglessness hasn't been resolved, only ignored.

In light of this analysis, I hope to have demonstrated the conventional arguments against self-killing as a means of preserving well-being, alongside their logical deficiencies per the three criticisms explicated upon throughout: firstly, that pragmatic optimism is not an option available to all, and fails in resolving the concerns leading to consideration of self-killing. Secondly, that life gives more reason to be pessimistic than optimistic, at least over the course of a single life. And thirdly, that attempts to undermine the profundity of existential concerns have been unsuccessful – that they recommend forging of subjective meaning where the subjectivity of meaning was the catalyst for angst, or a volitional blindness to existential concerns, the value of which was dismissed per my first argument. As stated in my preamble, these claims are not made lightly. Due to the unpopularity of pessimism within philosophy, progress within the school has been slow, and whilst there is worth in having an account such as this of what progress has been made within it, there is a much greater worth in presenting these arguments for the end of their revocation. It is my sincere hope that no one should read these words and not be highly sceptical.

Bibliography

Aronson, R. (Last viewed: 25/11/2017) Suicide, Absurdity and Happiness: The Myth of Sisyphus, *Stanford Encyclopaedia of Philosophy,* forthcoming URL = <https://plato.stanford.edu/entries/camus/#SuiAbsHapMytSis>

Benatar, D. (2017) *The Human Predicament*. Oxford University Press: New York.

Benatar, D. (2008) *Better Never to Have Been: The Harm of Coming into Existence*. Oxford University Press: New York.

Camus, A., O'Brien, J. (2005) *The Myth Of Sisyphus (Penguin Great Ideas)*. Penguin Group: London.

Cioran, E., Howard R. (2010) *A Short History of Decay (Central European Classics)* Penguin Group: London.

Hobbes, T., Gaskin, J. (2008) *Leviathan (Oxford World Classics)*. Oxford University Press: New York.

Miller, J. et al. (2015) *Utopic pessimism: The messianic underpinnings of the antinatalist polemic*. ProQuest Dissertations and Theses.

Nietzsche, F., Hollingdale, R. (1974) *Thus Spoke Zarathustra*. Penguin Group: London.

Nietzsche, F., Kaufmann, W. (1991) *The Gay Science*. Random House Inc.: USA.

Schopenhauer, A., Norman, J. (2010) *The World as Will and Representation, Volume 1*. Cambridge University Press: Cambridge.

Shade, P., & Lachs, John. (1997). *Habits of Hope: A Pragmatic Theory of the Life of Hope,* ProQuest Dissertations and Theses.

Weijers, D. (Last viewed 25/11/2017) Value and Prudential Hedonism, *Internet Encyclopaedia of Philosophy*, forthcoming URL = <https://www.iep.utm.edu/hedonism/#SH1b>

Wicks, R. (Last viewed: 25/11/2017) The World as Will, *Stanford Encyclopaedia of Philosophy,* forthcoming URL = < https://plato.stanford.edu/entries/schopenhauer/#4>

Does the distinction of temporal parts effectively differentiate the endurantist and perdurantist approaches to persistence?

Violet Tinnion

Both endurance theory and perdurance theory argue that objects can persist through time, though they offer different accounts of the way in which this happens. This essay will first discuss the differences between the two positions, arguing that the classical distinction of temporal parts is not sufficient to adequately capture their dissimilarity. I argue that the idea of extension over time allows us to draw a clearer distinction between the two approaches. Secondly, I shall more closely examine the enduransist approach to persistence, arguing that the main motivation behind this theory is that it is the "common-sense" view. I shall show that in order to solve the problem of temporary intrinsics, it must be made so complex that it undermines this initial motivation. Next I will discuss the worm-theory perdurantist approach, arguing that it offers us a more convincing answer to the problem of temporary intrinsics. Finally I will demonstrate that, although worm-theory is a more convincing approach to persistence than endurantism, there are alternative formulations of perdurantism that deserve consideration.

Endurantists and perdurantists agree that things can exist at more than one time, i.e. that they can persist. What they disagree on is how things persist. The classical distinction drawn between the two approaches is that perdurantists believe in temporal parts whilst endurantists do not. For the purpose of this essay, I will use Heller's definition of "temporal parts" which states that a temporal part, call it P, of object O is something which comes into existence at t1 and goes out of existence at t2, and takes up some portion of the space that O occupies for all the time that P exists (1983, p.323). Endurantists argue that at any point in time I am "wholly present". I move through time in my entirety with all of my parts present at any given moment

(Merricks, 1999, p.424). In contrast, perdurantists argue that my current phase is merely a temporal part of me, and that all of my parts are never fully present at one moment in time. For example, imagine that I attend a football match for five minutes and then leave. We would not say I have "seen" the match. I was only present for a short time and so only saw a part of it. Similarly, perdurantists argue that whenever we view an object we only see a temporal part of it rather than seeing the object in its entirety.

Whilst it is common for perdurantists to posit the existence of temporal parts and for endurantists to reject them, this idea is not sufficient to capture the dissimilarity between the two approaches. An endurantist may believe that there are "at least some enduring temporal parts", whilst maintaining that objects wholly exist through time (McKinnon, 2002, p.294). For example, I have lost teeth growing up and have gained new hairs. These are parts of me which lie in the past (teeth) and in the future (hair). According to Heller's definition above, my teeth and my hair are thus temporal parts of me. Moreover, a perdurantist may argue that it is possible for an object to perdure yet lack temporal parts (Merricks, p.431). Imagine an organism that is composed of four-dimensional cells, and these cells are the only proper parts that organism has. The organism thus lacks temporal parts, but we may still argue that it perdures given that it "has one more dimension than space" (*ibid*). Hence the idea of dimensions and extension through time allows us to draw a clearer distinction between the two approaches. Perdurantists argue that objects have extension over time in the same way that they have extension through space, i.e. they are four-dimensional (van Inwagen, 1990, p.245). For example, when looking at a giraffe we would say that its head is "up there" and its tail is "over there". It takes up space by having different parts in different places. Objects are stretched across time in the same way as the giraffe is stretched across space. Though it is often common for perdurantists to believe in temporal parts, this four-dimensionalist account of objects does not necessitate their existence. Endurance theorists reject this idea, maintaining that objects are three-dimensional and only have extension through space.

Having drawn a clearer distinction between the two approaches to persistence, I will now closely examine endurance theory with respect to the problem of temporary intrinsics. Imagine that I have a bottle of milk in my kitchen. On Monday the milk is sweet, and on Sunday the milk is sour. If the milk is "wholly present" throughout the week, it seems we are saying that it has both the intrinsic properties of "being sweet" and of "being sour" at the same time. However, this is contradictory as it does not seem possible that something can be totally sweet and totally sour at the same time. Solving this problem is important because we accept that persisting objects can hold contradictory properties and so we need a theory that is able to explain how this is the case (Lewis, 2001, p.203). The strongest formulation of endurance theory argues that objects hold "time-indexed relations" (van Inwagen, p.247). This formulation states that the milk does not have the conflicting intrinsic properties of "being sweet" and "being sour" at the same time, but instead bears the relation of "being sweet at" to Monday and "being sour at" to Sunday. The milk is still "wholly present", but is now able to hold conflicting properties because they "stand in different relations to different times" (Hawley, p.16).

I shall now consider a criticism to the above account. The three-dimensionalist endurance approach is often seen as the "common sense view" and the approach that simply explains what we see around us (Hawley, 2004, p.11). For example, when reading a book, it seems as though I am viewing the object in its entirety rather than a temporal part of it. Given that this is the main motivation for accepting the theory, we ought to evaluate its success based on whether it can fulfil this aim. On closer examination, however, endurance theory becomes less intuitive. If all the features of the milk are relational (e.g. being sweet, being cold, being liquid) then it seems we have reduced the milk to almost nothing because it now has very few (if any) intrinsic properties. For example, if the milk has different colour-relations to different times, then it seems to have no colour of its own. This "downgrades" objects to be "massless, colourless, shapeless and so on" (Hawley, 1998, p.213).

One could respond to this concern by arguing that seeing these features as relations is not the same as saying they do not have these features at all. The object has some intrinsic properties which determine the colour-relations, mass-relations and so on. However, this response is futile. Because most of the features of any object are relational, the objects have very few intrinsic properties. This means we cannot explain the differences between objects without making their temporary features essential (Hawley, 2004, p.19). For example, when explaining the difference between bread and chocolate, I would say that one is white and one is brown, that one is soft and one is hard and so on. Here I am referring to the essential intrinsic properties of the objects. Once we reduce these properties to relations, I have no way of explaining the difference between the two objects. This is an unappealing consequence which undermines the initial motivation behind endurance theory. Thus we must consider an alternative approach to persistence.

Perdurance theory avoids the problem outlined above and thus offers a more convincing approach to persistence. The classical formulation of perdurantism argues objects are space-time worms and persist through time because they are "composed of different temporal stages" (Lewis, p.41). For example, the milk in my kitchen has the intrinsic property of "being sweet" on Monday and a different intrinsic property of "being sour" on Sunday. It is able to change its intrinsic properties in such a way because it has temporal parts (*ibid,* p.203). So the temporal part "being sweet" comes into existence on Monday and goes out of existence on Sunday when a different temporal part ("being sour") comes into existence. The milk can hold both the contradictory properties of "being sweet" and "being sour" because it holds them at different times, and thus avoids the problem of temporary intrinsics. Moreover, we are able to describe the milk by referring to its many different temporal parts and thus avoid reducing the object to something that is "massless, colourless, shapeless" like the endurantist (Hawley, 1998, p.213). This is a four-dimensionalist view of objects because on Monday we are only seeing a particular phase of the milk (i.e. the "being sweet" phase) rather than seeing the milk in its entirety (Heller,

p.328). Because this view is able to avoid the problem of temporary intrinsics and of reducing objects to time-indexed relations, I argue it offers us a more convincing account of persistence.

Having explained how perdurance theory provides a successful solution to the problem of temporary intrinsics, I shall now consider a counter-argument. The four-dimensionalist view often faces criticism as it seems that when we view objects we see them in their entirety. When examining endurance theory, though, we have seen that this view quickly becomes counter-intuitive. Furthermore, I argue that the perdurantist account is not as perplexing as it initially seems to be. Endurance theory tells us that the milk exists on Monday and on Sunday, whereas perdurance theory tells us it exists from Monday until Sunday. We should think of objects as "existing within regions of time" (*ibid*, p.325). This is similar to the way we talk about events. For example, we say the birthday party was from noon until two, rather than at two specifically. Thus, we just need to adjust the way we speak about objects to mirror the way we speak about events. Given that we already use this kind of temporal language in our everyday conversation, it does not seem so counterintuitive as to warrant us rejecting perdurance theory. Thus I maintain that perdurance theory offers us a more successful approach to persistence when compared with endurance theory.

However, whilst this classical formulation of perdurantism offers us a more convincing approach to persistence than endurantism, there are other theories of perdurance that deserve consideration. Sider argues objects are not extended through time like a space-time worm, but they are extended with multiple distinct "stages" (2001, p.191). The milk is identical to its present stage. It is connected to its past and future stages which enables it to perdure through time. This avoids the problem of temporary intrinsics because the properties "being sweet" and "being sour" are held by different stages of the milk. Moreover, this form of perdurance is able to deal well with various puzzles that worm-theory fails to answer intuitively. Imagine that I exist at t1 and t2, but at t3 get divided in two people: A and B.

Worm-theory would say that, up until division, there are two overlapping space-time worms. However, if people are space-time worms, this implies that two people are occupying the same space at the same time (Sider, 1996, p.439). If I weigh 75kg and there are two worms overlapping, then there should be two masses of 75kg. This is evidently false. The stage view handles this problem better because it says that before division there is only one stage. If persons are identical to stages, then this means that there is only one person prior to division (*ibid,* p.441). This avoids the problem worm-theory ran into and is the more intuitive conclusion. Confusingly, Sider's view implies that, if a person is identical only to their present stage, then there have been as many people writing this paper in the last hour as there have been instants of time. Whilst Sider maintains these person-stages are psychologically connected with each other, this does seem like a counterintuitive consequence. Nevertheless, the stage view offers us strong reasons to consider alternative formulations of perdurance theory.

To conclude, this essay has argued that the classical distinction of temporal parts between endurantism and perdurantism is not sufficient to capture the difference between the two approaches. Instead, the idea of extension through time allows us to draw a sharp distinction between the theories. Whilst endurance theory is often seen as the "common sense view", when scrutinised it must resort to counter-intuitive claims to protect itself from criticism. Thus, we are motivated to consider other approaches to persistence. Perdurance theory offers us a more convincing account of persistence, providing a successful solution to the problem of temporary intrinsics and avoiding the problems endurance theory faces. Whilst the classical worm-theory formulation is the most commonly accepted, the stage view offers us an alternative form of perdurantism that deserves consideration.

Bibliography:

Hawley, K. (1998). Why Temporary Properties Are Not Relations between Physical Objects and Times. *Proceedings of the Aristotelian Society,* 98, pp.211-216.

Hawley, K. (2004). *How Things Persist.* Oxford: Clarendon, pp.9-36.

Heller, M. (1983). Temporal Parts of Four Dimensional Objects. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition,* 46(3), pp.323-334.

van Inwagen, P. (1990). Symposia Papers: Four-Dimensional Objects. *Noûs,* 24(2), pp.245-255.

Lewis, D. (2001). *On the Plurality of Worlds*. Malden, Mass; Oxford: Blackwell Publishers, pp.41-53 and pp.201-205.

McKinnon, N. (2002). The Endurance/Perdurance Distinction. *Australasian Journal of Philosophy,* 80(3), pp.288-306.

Merricks, T. (1999). Persistence, Parts and Presentism. *Noûs,* 33(3), pp.421-438.

Sider, T. (1996). All the world's a stage. *Australasian Journal of Philosophy,* 74(3), pp.4433-453.

Sider, T. (2001). *Four-Dimensionalism*. Oxford: Clarendon, pp.140-209.

Moral Realism and the Rejection of Normative Supervenience

Ben Jenkins

The supervenience of normative phenomena upon natural phenomena has posed a persistent problem for non-naturalist moral realism (henceforth 'Moral Realism'). I will consider a radical solution to this problem – the rejection of normative supervenience. This is a strategy Alison Hills (2009) adopts with her imaginative-resistance argument. In response to Hills' work, I will advance a conditional thesis: if normative supervenience can be *locally* formulated in a satisfactory manner, then Hills' rejection of normative supervenience (and her defence of Moral Realism) will fail. First, however, I will formulate a working definition of normative supervenience and present the 'supervenience problem' faced by Moral Realism.

'If it's wrong for them, why should it be right for you?' Such schoolyard sentiments echo through many of our childhoods, and the message remains salient in our day-to-day lives. The notion that these sentiments touch upon is the supervenience of normative phenomena upon relevant natural phenomena. Broadly speaking, X supervenes upon Y when, as a matter of metaphysical necessity, there can be no change in X without some change in Y. Our schoolyard misconduct, for instance, could not have been right for us whilst wrong for another because there is no relevant natural changes to underpin the normative change between our actions and the actions of others.

Supervenience relations are formulated with a supervening layer and a base layer. For us these are, respectively, normative and natural phenomena. Here natural phenomena can be understood as the kind of phenomena that could figure in a scientific theory (electrons, for example). Although this is a contentious definition of natural phenomena, the repercussions of such debates will be set aside for the purpose of this essay. I will formulate normative supervenience like so:

(S): *As a matter of metaphysical necessity, if two phenomena differ in their normative states they must differ in their natural states.*[31]

Whilst (S) may appear rather technical, it is simply a formulation of the schoolyard sentiment mentioned above. When two scenarios are identical in the natural sense (same actions, places $_{etc}$), (S) dictates that the normative nature of those two scenarios must also be identical. One scenario could not, for instance, be right whilst the other were wrong.

[31] For theoretical neutrality we ought to adopt the term 'non-normative/natural', as Ridge (2007) suggests. I will, however, stick to the term 'natural' for the sake of simplicity.

Let us consider how (S) challenges Moral Realism. Adopting realism about moral phenomena commits one to the existence of at least some mind-independent, normative phenomena. Non-naturalism then commits one to the notion that normative phenomena are distinct from, and *irreducible to*, natural phenomena[32]. Now, (S) involves a metaphysical necessity that ought to be explained – why, we might ask, does normative variance necessitate natural variance? This explanatory demand *is* the supervenience problem. The most obvious solution is that normative states are reducible to natural states, and so variation in the former necessitates variation in the latter. This solution, however, is reductionist so *non-naturalist* moral realism cannot accommodate it and remains particularly challenged by the supervenience problem.

In response to the supervenience problem, Moral Realists like Shafer-Landau (2009) typically endeavour to address the explanatory challenge of (S) whilst avoiding reductionism. Challengers then try to show that these explanations appeal to brute facts, or harbour some otherwise undesirable feature. Alison Hills, however, argues that we ought to reject normative supervenience altogether. Hills' argument requires at least two conditions. First, Hills needs some alternative to (S) which avoids the explanatory challenge, whilst explaining the relationship between normative and natural phenomena; second, Hills needs to show that it is possible to reject (S).

Hills' alternative is a much weaker claim than (S). She argues that the covariance we expect between normative and natural phenomena is not the result of any metaphysical necessity. Rather, according to Hills, such states exhibit a Constant Conjunction; that is, "[…] in the actual world, there are no differences in moral properties without differences in (some interesting subset of the) natural properties." (2009, pp.167). This account involves no metaphysical necessities and so avoids the explanatory demand of (S), potentially solving the supervenience problem.

Whilst Hills' Constant Conjunction is a controversial thesis, I will assume that it is a satisfactory alternative to (S). Instead I will focus on the second condition Hills must establish: that (S) can be rejected in the first place. I will begin by examining what reasons there are for positing normative supervenience, then challenge Hills' rejection of these reasons and consequently her defence of Moral Realism.

A prospective motivation for positing (S) is that it is conceptually necessary. Ridge (2007) provides support for this by taking something immoral, the acts and motives of Hitler, and arguing that no competent user of normative concepts can conceive of these natural states being moral, unless some natural change were to occur. The inconceivability of ¬(S) vis. '*not* (S)', illustrated by Ridge's example, is then taken as grounds for the conceptual necessity of (S).

Hills (2009) charges Ridge's argument with a suppressed premise. She argues that Ridge's argument requires the following: if we are unable to conceive of ¬(S) it is *because* (S) is inconceivable. Hills rejects this premise, arguing that the reason we cannot conceive of ¬(S) is that we experience imaginative resistance to worlds in which (S) is false. This avoids the actual impossibility of worlds in which (S) is false

[32] Non-naturalism also opposes supernatural normative phenomena, though the significance of supernatural phenomena will not be covered in this essay.

for the simple reason that "[…] classic imaginative resistance arises when *we can't because we won't."* (Gendler, 2006, pp.164).

Hills makes two moves to establish imaginative resistance in this context. First she highlights reasons that we don't want to consider normative variation, especially when a strong good-to-bad variation occurs - "[…] we think it is somehow contaminating […] or perhaps habit-forming […]" (2009, pp.172). Then she argues that these desires, however weak, explain the inconceivability of ¬(S). This is done by demonstrating that our imaginative resistance to conceivability cases is reduced both when the magnitude of normative variation is lowered and when we consider good-to-bad variations. It appears, for example, that it is easier to conceive of the normative inversion of Barack Obama's thoughts and actions, than it is in Hitler's case. Implementing imaginative resistance, like so, enables Hills to reject (S) and solve the supervenience problem.

To re-establish the supervenience problem we have to re-formulate (S) so that ¬(S) really is inconceivable. Supervenience comes in local and global versions. Global supervenience takes natural bases as entire possible worlds, whereas local supervenience takes them as natural states within a world. (Global-S) entails that full descriptions of a world *could not* have realized different normative states. Local supervenience, in contrast, only opposes worlds in which natural states could realise different normative states *at different times*. This makes (Local-S) a weaker claim; this means that its negation, '¬(Local-S)', has fewer conditions for us to conceive of, making it a stronger basis than (Global-S) for the conceptual necessity of (S).

I will now argue that we can use (Local-S) to make ¬(S) truly inconceivable. To do this I will form a new argument using a premise similar to Moore's Open Question Criterion (Moore, 1903). The plausible criterion is this:

If no fully comprehending individual is capable of asking a sincere question as to the necessity of a feature for a given concept, then that feature is conceptually necessitated by that concept.

Given the plausible criterion, the brunt of my argument will depend upon the next step - demonstrating that (Local-S) passes the criterion in relation to normative phenomena. Turning to this, consider the following scenario: take an intelligent person. Let us say she is a CEO. One year, having had their tax returns inspected, the CEO's company is fined for tax evasion. She admits guilt, is publicly defamed for her normative digressions and exhibits genuine regret. A year passes, nothing changes and the CEO finds herself in the exact same position. She then asks her secretary "is tax evasion still immoral?" Now, our CEO could be inquiring into the legality of her actions, or possibly public opinion. I suggest, however, that she could not be sincerely inquiring as to the normative nature of the offence, *the second time round*. This means that (Local-S) is necessitated by normative concepts, cannot be rejected, and remains a problem for Moral Realism.

The Moral Realist is faced with the challenge of finding a counter-example to the CEO-case. They may still feel, though, that we need to link inability to inconceivability. I think this is somewhat true; my argument does only establish our inability to conceive of a person that fits the plausible criterion. However, the use of local supervenience, and my form of argument, now avoids our aversion to

alternative moral states and our reluctance to consider them – the main motives Hills wielded as alternatives to the impossibility of ¬(S). This is done by asking the Moral Realist to consider a *transition* between two normative states within a given world, as opposed to different normative states simpliciter. Consequently, our inability to conceive of counter examples to (Local S) acts as evidence for the conceptual necessity of local normative supervenience. To see what has been achieved, consider the force of our argument in the cases where Ridge's argument seemed weak – namely, mild good-to-bad normative variations. It remains impossible, for example, that our CEO could be applauded for charitable donations one year, acknowledge the righteousness of her actions, and then sincerely ask whether charity is still good in the following year.[33]

I will now make two final comments on the thesis I've advanced. Firstly, I have presumed that (Local-S) can be formulated satisfactorily, but oversensitivity is a big obstacle to this. That is, if the bare individuality of two phenomena constitutes variation in the base, then (S) will be true in a trivial sense – because no two natural states *can* be the same. Though it is not possible to address this issue here, preventing the triviality of (Local-S) is essential for the application of my conditional thesis. Lastly, this essay has critiqued one approach to rejecting supervenience, but there may be more successful alternatives. Hopefully, however, the well-established use of open-questions to analyse conceptual relations in the Moral Realist literature will ensure that my argument has some traction with these alternatives.

This essay has explored reasons to think that normative supervenience might be rejected and consequently that the supervenience problem could be solved. It has provided a heuristic for establishing the conceptual necessity of one feature for a given concept and shown how this could be used to support a satisfactory, local formulation of normative supervenience. It is hence suggested that if the supervenience problem is to be solved, the Moral Realist will need to address the explanatory demand of normative supervenience, not reject it.

Bibliography

Blackburn, S. W. (1984). Supervenience Revisited. In I. Hacking (Ed.), *Exercises in Analysis* (pp. 59-74). Cambridge University Press. (Original work published 1981)

Hills, A. (2009). Supervenience and Moral Realism. In A. Hieke & H. Leitgeb (Eds.), *Reduction, Abstraction,* Analysis (pp. 163-177). De Gruyter.

Mackie, J. L. (1983). The Subjectivity of Values. In *Ethics: inventing right and wrong* (pp. 15-49). Harmondsworth, Middlesex, Eng.: Penguin Books.

Moore, G. E., & Moore, G. E. (1903). The Subject-Matter of Ethics. In *Principia Ethica* (pp. 1-36). Cambridge University Press. Retrieved October 26, 2017, from

[33] Here I presume that no new information has been uncovered on the matter of charitable donations. A re-estimation of the natural phenomena (for instance a discovery that charitable donations are, practically speaking, bad for the recipients of the CEO'S charity) could, of course, allow for a re-estimation of their normative nature.

https://www.learn.ed.ac.uk/bbcswebdav/pid-2638908-dt-content-rid-4994986_1/courses/PHIL100192017-8SV1SEM1/GE%20Moore.pdf.

Ridge, M. (2007) Anti-Reductionism and Supervenience. *Ethics and Moral Philosophy*, 147-169. doi:10.1163/9789004215337_010

Shafer-Landau, R. (2009). *Moral realism: A Defence*. Oxford: Oxford University Press.

Toppinen, T. (2017). Essentially Grounded Non-Naturalism and Normative Supervenience. *Topoi*. doi:10.1007/s11245-017-9456-x

Väyrynen, P. (2018). The Supervenience Challenge to Non-Naturalism. In T. C. McPherson & D. Plunkett (Eds.), *The Routledge handbook of metaethics* (pp. 170-184). Routledge, Taylor & Francis Group.