**Investigating Non-philosophers’ Judgements about the Asymmetry of Metaphysical Explanation**

**Abstract**

It is often supposed that metaphysical explanation is asymmetric: that for all x and y, if x metaphysically explains y, then y does not metaphysically explain x. Even amongst those who hold that metaphysical explanation is not asymmetric, but nonsymmetric, it is assumed that a relatively small number of particular explanations are symmetric: by and large, if x metaphysically explains y, then y does not metaphysically explain x. Both parties agree that as a matter of fact we at least typically judge that if x metaphysically explains y, then y does not metaphysically explain x. There has, however, been no empirical investigation of our judgements on these matters, beyond the armchair reflections of philosophers. This paper investigates our judgements and finds that in every case participants were presented with, a majority judged that x metaphysically explained y, and that y metaphysically explained x. This gives us reason to conclude not only that metaphysical explanation is not asymmetric, but also that it has many more symmetric instances than philosophers have supposed. Indeed, the evidence we collected is consistent with metaphysical explanations being symmetric. We evaluate the upshots of this research for theorising about metaphysical explanation.

**1 Introduction**

In what follows we suppose that metaphysical explanations are true propositions of the form ⌜x *because* y⌝[[1]](#footnote-1) where ⌜x⌝ and ⌜y⌝ are sentences, and *‘because’* expresses a particular kind of explanatory connection (i.e. one that is metaphysical rather than causal).[[2]](#footnote-2) It is often supposed that metaphysical explanations are transitive, asymmetric, irreflexive, and non-monotonic.[[3]](#footnote-3) Of late, these assumptions about asymmetry, transitivity and irreflexivity have come under fire, with some philosophers denying some, or all, of these claims.[[4]](#footnote-4) This paper focuses on just one such claim: that metaphysical explanation is asymmetric.

Technically, properties like asymmetry are properties of relations not of propositions. Still, we will say that metaphysical explanation is asymmetric just in case for any proposition of the form ⌜x *because* y⌝, p *because* q, if p *because* q is true, then q *because* p is not. Metaphysical explanation is symmetric just in case for any proposition of the form ⌜x *because* y⌝, p *because* q, if p *because* q is true then so is q *because* p. We will then talk of *particular* metaphysical explanations, such as p *because* q, being asymmetric, or symmetric. We will say that p *because* q is asymmetric just in case p *because* q is true, and q *because* p is false, and is symmetric just in case both p *because* q and q *because* p are true. Then metaphysical explanation is nonsymmetric just in case some propositions of the form ⌜x *because* y⌝ are symmetric, and some are asymmetric.

The idea that metaphysical explanation is asymmetric is motivated by the thought that circular explanations are not explanations at all. We find it deeply unsatisfying to be told both that p explains q, *and* that q explains p. Moreover, on a case-by-case basis we judge either that p explains q, or that q explains p, but not both. Recently, both these motivations have come under scrutiny.

Some have challenged the idea that there is something problematic about circles of explanation.[[5]](#footnote-5) If a explains b, and b explains c and c explains d, and d explains a, there need not, it is argued, be anything vicious, or otherwise concerning, about this circle of explanation. Still others have provided examples of what are argued to be particular symmetrical metaphysical explanations.[[6]](#footnote-6) For instance, Rodriguez-Pereya (2015) argues that truthmaking (which he takes to be a species of metaphysical explanation) is not asymmetric. Consider proposition A: <A exists>. That proposition says of itself that it exists. If true, then, it is made true by the existence of that very proposition. So this proposition is made true by itself. Hence there are reflexive metaphysical explanations and, it follows, there are symmetric metaphysical explanations since asymmetry entails irreflexivity.

A related example comes from Thompson’s (2016) truth-telling pair. P says that ‘Q’ is true, while Q says that ‘P’ is true. This, too, appears to be a case in which P explains Q, and Q explains P.

These examples of symmetry are recherché to say the least. So even if they show that metaphysical explanation is non-symmetric, they suggest (or at least are consistent with) the contention that most token metaphysical explanations are asymmetric: mostly, if p *because* q is true, then q *because* p is not. Indeed, to our knowledge no one has argued that metaphysical explanations are often, or always, symmetric. Rather, philosophers suppose that, at the very least, most metaphysical explanations are asymmetric.

So far, attempts to determine the extent of symmetrical metaphysical explanations (if any) have proceeded by consultation of philosophers’ own intuitions either about explanation in general (i.e. whether it can be circular) and by reflection on philosophical judgements about a range of particular metaphysical explanations. So, for instance, paradigm cases of metaphysical explanation strike philosophers as asymmetric. It is thought that the existence of some object explains the existence of the singleton set containing that object, that the fact that the ball is maroon explains why it is red, that the fact that action A maximising utility explains why it is right, that the fact that someone is in a certain brain state explains why they are in a certain mental state, and so on. In all of these cases, and more, philosophical intuition is that explanation proceeds in one direction only.

While reflection on our own judgements strikes us as a good place to start, it surely ought not be the final word on the matter.

After all, many philosophers suppose that metaphysical explanations have been a staple of human interaction for a long time, and are not a unique feature of philosophical discourse. Schaffer (2009:375) claims that metaphysical explanation is “a natural and intuitive notion, for which there exist clear examples, and clear formal constraints” and Dasgupta (2017:74-76) notes that metaphysical explanation is “intuitive and familiar…examples are ubiquitous. Why is a faculty meeting occurring? Because the faculty are gathered in a room discussing matters of importance to the department, etc. Why is this water hot? Because its mean kinetic energy is high. Why have I lost this game of chess? Because my king is in check-mate.” Indeed, it is “*an everyday concept used by the masses*. When I explain the concept to non-philosophers they recognize it immediately and talk intelligibly about it, offering examples of [metaphysical] explanations in their own fields of biology, economics, journalism, or cooking. To them it is not a new concept.” Further, Glazier (2020:121) notes that “[f]or its enthusiasts, [metaphysical] explanation is both ubiquitous in ordinary life and central to many of philosophy’s biggest questions”.

We take seriously the idea that the phenomenon under investigation—metaphysical explanation—is not some specialised philosophical notion with which the folk are unfamiliar. Rather, at least for now, we suppose that it is a phenomenon with which we all engage, and hence that the judgements of non-philosophers ought be taken into account when theorising about the phenomenon. Exactly what role these folk judgements ought play in our theorising is, we take it, very much open, and something to which we return in §4.

We begin, in §2, by outlining relevant empirical research and developing some hypotheses. We then describe the study’s methodology and results in §3. Our results show that non-philosophers judge a range of cases—including cases that philosophers judge to be cases of asymmetric explanation—to be ones in which the metaphysical explanations in question are symmetric. We consider the upshot of these results in §4.

**2 Empirical Evidence and Hypotheses**

There is little empirical work on judgements about metaphysical explanation,[[7]](#footnote-7) and no work on whether people judge particular metaphysical explanations to be symmetrical or asymmetrical. There is, however, a good deal of empirical work on explanation more generally, and, in particular, on causal explanation. That work is important in developing testable hypotheses about metaphysical explanation.

That research suggests that there is an intimate relationship between causal explanations, the imagining of interventions and the representation of interventional affordances (that is, representations that intervening on one part of the world will thereby intervene on some other part of the world). There is ample evidence that imagining interventions is vital in causal reasoning[[8]](#footnote-8) and in causal explanation.[[9]](#footnote-9) People tend to judge that [x] is a causal explanation for [y] only when they judge that intervening on [x] is a way to intervene on [y], but that intervening on [y] is not a way to intervene on [x], and when they judge that [x] is earlier than [y] (Sloman 2005).

In turn, research shows that people are inclined to prefer the explanation that is most salient at the context (Hilton and Slugoski 1986; Knobe 2009; Hitchcock and Knobe 2009; Kahneman & Miller 1986).

Sometimes a causal explanation is more salient when it appeals to causal factors that are in some way *abnormal*: to causal factors that are not typical and hence which stand out. Sometimes salience is the product of *practicality*. People are more likely to prefer causal explanations that appeal to causal factors over which they have control, to those that appeal to causal factors over which they have no control (Hitchcock and Knobe 2009). Further, different kinds of explanations are evaluated according to their perceived *usefulness* to the evaluator in facilitating the performance of upcoming tasks (Vasilyeva, Wilkenfeld, and Lombrozo 2017) and are sensitive to which contrast class is identified (McGill 1989; Hilton & Erb 1996. Sometimes salience is the product of *background beliefs and expectations*. People are more likely to prefer causal explanations when those explanations accord with their expectations about what explains what (Hitchcock and Kobe 2009; Kahneman and Miller 1986) and cohere with their existing set of explanations (Murphy & Medin, 1985; Mackonis, 2013; Chapman & Chapman (1969).

Returning to metaphysical explanation, we might expect to find similar connections between our judgements regarding metaphysical explanations, and these sorts of factors. Indeed, the idea that there is some connection between interventions and metaphysical explanation is not new. Miller & Norton (2017) make this case, while Schaffer (2016) and Wilson (2018a; 2018b) have independently developed accounts of grounding that make use of the same formalism as interventionist accounts of causation.

If we take these connections seriously, then, this research suggests that when people are asked to judge whether a proposition of the form ⌜x *because* y⌝ is an explanation, they will be more inclined to do so when, at the context they are judging, intervening on [y] in order to intervene on [x] is salient. Latham and Miller’s (ms) recent empirical work supports this hypothesis.

This suggests that if we describe a scenario in which intervening on [y] in order to intervene on [x] is salient, and another in which intervening on [x] in order to intervene on [y] is salient, we might find that people judge that ⌜x *because* y⌝is an explanation in the former condition, and that ⌜y *because* x⌝ is an explanation in the latter condition. Where people judge that a particular instance of ⌜x *because* y⌝, p *because* q, is an explanation in one condition, and that q *because* p an explanation in another condition, we will say that with they judge this explanation to be symmetrical, or that they have symmetrical judgements about this explanation.

We know that one way to make an intervention salient at some context is to make it practical and useful. So if we describe a circumstance in which intervening on [y] to intervene on [x] is practical and useful, then that intervention ought be salient in that condition, and *mutatis mutandis* if we describe a circumstance in which intervening on [x] in order to intervene on [y] is practical and useful. Given the connection between the salience of interventions and people’s judgements about causal explanations, we have reason to expect people to judge that ⌜x *because* y⌝ is an explanation in a condition in which intervening on [y] to intervene on [x] is salient, and to judge that ⌜y *because* x⌝ is an explanation in a condition in which intervening on [x] to intervene on [y] is salient. Hence we predict that people will have symmetrical judgements across those two conditions.

In what follows we focus on three cases of metaphysical explanation. Since familiarity with certain kinds of explanations (or with explanans and explanandum) might play a role in non-philosophers’ judgments, we chose cases that vary along this dimension.

The first, mind/brain case, is one that is relatively familiar to non-philosophers. It is also one in which philosophers have asymmetric judgments (judging that the properties of brains explain the properties of minds and not the other way around). In one vignette intervening on the brain in order to intervene on the mind is made salient, by making an intervention on the brain practical and useful through the provision of drugs. In the other vignette intervening on the mind in order to intervene on the brain is made salient, by making an intervention on the mind practical and useful through the provision of therapy.

The second case is the set/singleton case. This was chosen to be a case that is unfamiliar to non-philosophers. Again, this is a case in which philosophers have asymmetric judgements: judging that the existence of the member of the singleton set explains the existence of the set. In order to make the relevant interventions salient by being practical and useful, we described vignettes in which the individuals have different capacities than any of us: they are either able to directly intervene on abstract objects, including sets, or to travel anywhere in space and time and remove concrete objects from the timeline.

Finally, our third case is the Euthyphro case, in which even philosophers do not always agree about which direction explanation proceeds, although, notably, each philosopher tends to think it proceeds in only one direction. One vignette described a condition in which it was practical and useful to intervene on what God approves of, and the other described a condition in which was practical and useful to intervene on what is good.

We predicted that across all three kinds of cases people’s judgements about what explains what, would align with the intervention that was salient. In order to test this hypothesis in each condition we asked participants both whether ⌜x *because* y⌝ is true, and whether it is an explanation for the (relevant) subject.

In all, we had nine specific hypotheses. First, we hypothesised that people would, on average, agree that ⌜x *because* y⌝ is true at a condition in which intervening on [y] in order to intervene on [x] is made salient. That is, we hypothesised that mean levels of agreement that ⌜x *because* y⌝ is true in that condition would be significantly above 4, the mid-way point on a Likert scale that ran from 1 (strongly disagree) to 7 (strongly agree). Second, we hypothesised that a majority of participants would agree that ⌜x *because* y⌝ is true in this condition. Third, we hypothesised that mean levels of agreement that ⌜y *because* x⌝ is true, in a condition in which intervening on [x] in order to intervene on [y] is made salient, would be significantly above 4, and (fourth) that a majority of participants would agree that ⌜y *because* x⌝ is true in that condition.

Fifth, we hypothesised that mean levels of agreement would be significantly above 4 that ⌜x *because* y⌝ was an explanation *for the relevant subject,* of why x, in a condition in which intervening on [y] in order to intervene on [x] is made salient. Six, we hypothesised that a majority of participants would, in this condition, judge that ⌜x *because* y⌝ is an explanation for the relevant subject, of why x. Seventh, we hypothesised that mean levels of agreement would be significantly above 4 that ⌜y *because* x⌝ is an explanation *for the relevant subject,* of why y, in a condition in which intervening on [x] in order to intervene on [y] is made salient. Eighth, we hypothesised that a majority of participants would agree, in that condition, that ⌜y *because* x⌝ is an explanation for the relevant subject of why y.

Ninth, and finally, we hypothesised that there would be no significant difference between participant’s levels of agreement regarding whether ⌜x *because* y⌝ is true, and their levels of agreement regarding whether ⌜x *because* y⌝ is an explanation for the relevant subject.

**3 Experimental Design**

**3.1 Method**

*3.1.1 Participants*

625 people participated in the study. Participants were U.S. residents, recruited and tested online using Amazon Mechanical Turk,[[10]](#footnote-10) and compensated $2 for approximately 20 minutes of their time. 65 participants had to be excluded for failing to follow task instructions. This means that they failed to answer the questions (89), or failed an attentional check question (115). The remaining sample was composed of 421 participants (aged 21-99; 155 female, 2 transgender/non-binary). Mean age 32.07 (SD = 10.07). Ethics approval for this study was obtained from the [blanked] Human Research Ethics Committee. Informed consent was obtained from all participants prior to testing. The survey was conducted online using Qualtrics.

*3.1.2 Materials and Procedure*

Participants were divided into six groups, each of which saw a single vignette and responded to one set of statements. Group 1 participants saw vignette 1(a), and group 2 participants saw vignette 1(b).

Vignette 1(a) Mind/Brain

Fred is experiencing a lot of stress. He learns that whenever someone is stressed, they have very high cortisol levels, and that whenever anyone has very high cortisol levels, they are stressed. He learns that stress and high cortisol levels go together because what it is to have one, is to have the other.

Fred goes to his pharmacist, Maria. Maria tells Fred that a new drug has just been released—*destressor*—and that this drug works by preventing cortisol from being released into the system. Maria tells Fred ‘Fred, you are stressed because you have high cortisol levels’ and she recommends that Fred takes destressor.

Vignette 1(b) Mind/Brain

Fred is experiencing a lot of stress. He learns that whenever someone is stressed, they have very high cortisol levels, and that whenever anyone has very high cortisol levels, they are stressed. He learns that stress and high cortisol levels go together because what it is to have one, is to have the other.

Fred goes to his therapist, Maria. Maria tells Fred that a new therapy technique has been developed—*unstressor*—and that this therapy, consisting of meditation and mindfulness, works to alleviate the experience of stress. The therapist tells Fred ‘Fred, you have high cortisol levels because you are stressed’ and she recommends that Fred undergo the unstressor therapy.

Group 3 participants saw vignette 2(a) and group 4 participants saw vignette 2(b)

Vignette 2a Sets

Sets are abstract objects, like numbers, or functions (like addition, and subtraction, and division). But unlike number or functions, sets have members. Even though sets are abstract objects, their members can be concrete things, like dogs, chairs, electrons, and people. In fact, for any bunch of things, there’s a set containing just those things. *Singleton* sets are sets that contain only one member. So the singleton set containing the number 2, is the set that only contains 2. The singleton set containing Eminem, is the set that contains only Eminem as a member. Sets only exist when their members do, and their members only exist when the set does. So if Eminem exists, then the singleton set containing Eminem exists. If the singleton set containing Eminem exists, then Eminem exists. If one exists, then so does the other.

Fred is a member of an elite secret time team. This team of people can travel anywhere in time, and can erase people and events from the timeline, so that those people and events never existed at all. Fred is talking with his supervisor, Maria, about the suggestion that they ought to erase Eminem from the timeline. Fred says to Maria ‘but what about the singleton set containing Eminem?’ Maria responds that if they erase Eminem, then they will erase the singleton set as well. At the end of their discussion, Maria tells Fred that the singleton set containing Eminem exists because Eminem exists.

Vignette 2b Sets

Sets are abstract objects, like numbers, or functions (like addition, and subtraction, and division). But unlike number or functions, sets have members. Even though sets are abstract objects, their members can be concrete things, like dogs, chairs, electrons, and people. In fact, for any bunch of things, there’s a set containing just those things. *Singleton* sets are sets that contain only one member. So the singleton set containing the number 2, is the set that only contains 2. The singleton set containing Eminem, is the set that contains only Eminem as a member. Sets only exist when their members do, and their members only exist when the set does. So if Eminem exists, then the singleton set containing Eminem exists. If the singleton set containing Eminem exists, then Eminem exists. If one exists, then so does the other.

Fred is a member of an elite secret abstract-objects team. This team of people can travel anywhere in the abstract realm and can intervene on any abstract object. The team of people can, for instance, intervene on numbers, and functions and sets. Fred is talking with his supervisor, Maria, about the suggestion that they ought to erase the singleton set containing Eminem, from the abstract realm. Fred says to Maria ‘but what about Eminem?’ Maria responds that if they erase the singleton set containing Eminem, from the abstract realm, then they will erase Eminem as well. At the end of their discussion, Maria tells Fred that Eminem exists because the singleton set containing Eminem exists.

Group 5 participants saw vignette 3(a) and group 6 saw vignette 3(b).

Vignette 3a Euthyphro

Fred is not sure whether or not God approves of eating snails. Fred finds snails really tasty, and hopes that God does approve of eating snails, since, Fred knows, God approves of things only when they are good, and everything that is good, is something God approves of.

One day Fred meets up with his old friend, Maria, and the two of them visit Maria’s prayer group: the Change God’s Mind prayer group. The group aims to pray to God, to change His mind about what he approves of. Maria suggests that Fred prays, with the group, to bring it about that God approves of eating snails. In fact, God exists, and He does approve of eating snails. After the prayer meeting, Maria tells Fred that eating snails is good because God approves of it.

Vignette 3b Euthyphro

Fred is not sure whether or not God approves of eating snails. Fred finds snails really tasty, and hopes that God does approve of eating snails, since, Fred knows, God approves of things only when they are good, and everything that is good, is something God approves of.

One day Fred meets up with his old friend, Maria, and the two of them visit Maria’s prayer group: the Change What is Good prayer group. The group aims to pray to God, to get God to change which things are good. Maria suggests that Fred prays, with the group, to bring it about that eating snails is good. In fact, eating snails is good, and God exists, and He does approve of eating snails. After the prayer meeting, Maria tells Fred that God approves of eating snails because eating snails is good.

After seeing one of these vignettes all participants were asked to respond to Maria’s assertion on two different Likert scales. One of the Likert scales ran from 1 ‘Completely sure that what Maria says is false’ at one end (either the far left or the far right, determined randomly: this is so on all the Likert scales henceforth described) to 7 ‘Completely sure that what Maria says is true’ at the opposite end of the scale via 4 ‘I am indifferent between these two options’. The other Likert scale ran from 1 ‘Completely sure that the statement is *not* an explanation for Maria’ at one end to 7 ‘Completely sure that the statement *is* an explanation for Maria via 4 ‘I am indifferent between these two options’.

After having done so, participants were taken to a new page that did not have either the vignette or Likert scales on it and were asked an attentional check question: *“In the vignette you were asked to read, what were Fred and Maria talking about?”* to which they could answer (1) Abstract Objects; (2) Bicycles; (3) God or (4) Minds and Brains. Participants who choose incorrectly were eliminated.

*3.1.3 Analyses*

Call participants’ levels of agreement regarding whether Maria’s assertion is true, their *levels of truth agreement.* Then participants whose level of truth agreement is *higher* are participants who more strongly agree that Maria’s assertion is true. Call participants’ level of agreement regarding whether Maria’s assertion is an explanation for Maria, their *levels of explanation agreement*. Then participants whose level of explanation agreement is higher are participants who more strongly agree that Maria’s assertion is an explanation for her.

In order to test whether people’s levels of truth agreement and levels of explanation agreement differed significantly from indifference (a score of 4 on the Likert scale) we ran separate one-sample t-tests to test whether the mean response significantly differs from 4 in each condition. If the mean is significantly above 4, then overall people might think that what Maria says is true or is an explanation for her; if the mean is significantly below 4 then overall people might think that what Maria says is false or is *not* an explanation for her; if the mean does not differ significantly from 4 then overall people might be indifferent. We compared levels of truth agreement and levels of explanation agreement between conditions 1(a) and 1(b), 2(a) and 2(b), and 3(a) and 3(b) using separate between-subjects t-tests.

Since mean results can be misleading (a mean of 5 can be the result of a smaller number of people choosing 7, and a larger number choosing 3) we also want to know whether a majority of people agree, or disagree, that a certain proposition is true, or is an explanation for Maria.

In the conditions in which the mean was significantly greater than 4, we combined the proportion of people who thought that what Maria said was false, or was not an explanation, with those who were indifferent. We then ran separate one-way 𝜒2-tests to test whether the *majority* of people responded in agreement that what Maria said was true, or was an explanation for her. We compared the proportions of people between conditions 1(a) and 1(b), 2(a) and 2(b), and 3(a) and 3(b) using separate 𝜒2 tests of independence.

Finally, in order to test whether within participants their levels of truth agreement differed significantly from their levels of explanation agreement we ran separate paired-sample t-tests for each condition.

**4 Results**

Consider, first, our two hypotheses about people’s judgements about the truth of the various propositions. We hypothesised that in a condition in which intervening on [y] in order to intervene on [x] is made salient, participants’ mean levels of truth agreement that ⌜x *because* y⌝ is true would be significantly above 4, and that in a condition in which intervening on [x] in order to intervene on [y] is made salient, participants’ mean levels of agreement that ⌜y *because* x⌝ is true would be significantly above 4.

Table 1 below summarises the descriptive data that speaks to these hypotheses.

The ‘Yes’ column represents the proportion of participants who reported that what Maria said is true (5, 6 or 7). The ‘No’ column represents the proportion of participants who reported that what Maria said is false (1, 2 or 3). The ‘I’ column represents the proportion of people who reported being indifferent between these two options (4). The t-value tells us the result of the t-test, and the p-value tells us whether the t-test is significant. Where the p-value is <0.05, the t-test tells us that the mean is significantly above, or below, 4.

As we can see, both hypotheses were vindicated across *all three* pairs of vignettes.

*Table 1. Descriptive data and t-test results for level of truth agreement.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Condition** | **%Yes** | **%No** | **%I** | **Mean** | **SD** | ***t-value*** | ***p*-value** |
| **Mind/Brain** | | | | | | | |
| **1A:**  **Stress because cortisol** (N = 74) | 71.6 | 16.2 | 12.2 | 5.09 | 1.48 | 6.353 | <.001 |
| **1B: Cortisol because stress** (N = 71) | 77.5 | 7.0 | 15.5 | 5.46 | 1.34 | 9.212 | <.001 |
| **Between-Subjects:** |  | | | | | -1.576 | .117 |
| **Sets** | | | | | | | |
| **2A: Set because member** (N = 56) | 85.9 | 5.4 | 8.9 | 5.66 | 1.33 | 9.380 | <.001 |
| **2B: Member because set** (N = 64) | 71.8 | 14.1 | 14.1 | 5.17 | 1.60 | 5.862 | <.001 |
| **Between-Subjects:** |  | | | | | 1.808 | .073 |
| **Euthyphro** | | | | | | | |
| **3A: Good because God approves** (N = 76) | 65.8 | 15.8 | 18.4 | 4.96 | 1.66 | 5.042 | <.001 |
| **3B: God approves because good** (N = 80) | 67.5 | 17.5 | 15.0 | 4.86 | 1.63 | 4.737 | <.001 |
| **Between-Subjects:** |  | | | | | .372 | .710 |

Next, consider our two hypotheses regarding people’s judgements about whether ⌜x *because* y⌝ is an explanation for the relevant subject. We hypothesised that mean levels of agreement that ⌜x *because* y⌝ is an explanation for the relevant subject of why x, would be significantly above 4 in a condition in which intervening on [y] in order to intervene on [x] is made salient. We hypothesised that mean levels of agreement that ⌜y *because* x⌝ is an explanation for the relevant subject of why y, would be significantly above 4 in a condition in which intervening on [x] in order to intervene on [y] is made salient.

Table 2 below summarises the descriptive data that speaks to these hypotheses.

The ‘Yes’ column represents the proportion of participants who reported that what Maria said is an explanation for her (5, 6 or 7). The ‘No’ column represents the proportion of participants who reported that what Maria said is not an explanation for her (1, 2 or 3). The ‘I’ column represents the proportion of people who reported being indifferent between these two options (4).

*Table 2. Descriptive data and t-test results for level of explanation agreement.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Condition** | **%Yes** | **%No** | **%I** | **Mean** | **SD** | ***t-value*** | ***p*-value** |
| **Mind/Brain** | | | | | | | |
| **1A:**  **Stress because cortisol** (N = 74) | 71.6 | 16.2 | 12.2 | 5.00 | 1.54 | 5.604 | <.001 |
| **1B: Cortisol because stress** (N = 71) | 81.6 | 8.5 | 9.9 | 5.28 | 1.35 | 7.975 | <.001 |
| **Between-Subjects:** |  | | | | | -1.170 | .244 |
| **Sets** | | | | | | | |
| **2A: Set because member** (N = 56) | 78.6 | 12.5 | 8.9 | 5.45 | 1.50 | 7.213 | <.001 |
| **2B: Member because set** (N = 64) | 75 | 17.2 | 7.8 | 4.95 | 1.58 | 4.832 | <.001 |
| **Between-Subjects:** |  | | | | | 1.748 | .083 |
| **Euthyphro** | | | | | | | |
| **3A: Good because God approves** (N = 76) | 68.4 | 15.8 | 15.8 | 4.87 | 1.58 | 4.799 | <.001 |
| **3B: God approves because good** (N = 80) | 73.7 | 16.3 | 10.0 | 4.88 | 1.62 | 4.838 | <.001 |
| **Between-Subjects:** |  | | | | | -.026 | .980 |

Table 2 shows that both hypotheses are supported. The results of our one-sample t-tests (which, recall, show us whether the mean response differs significantly from a value of 4) appear to show that overall, people think that what Maria says is an explanation for her, in *all* the conditions we tested.

However, the one-sample t-test’s do not tell us what the majority of participants judge in any condition. For that we must look to the results of our one-way 𝜒2-tests.

We hypothesised that a majority of people would judge that a condition in which intervening on [y] in order to intervene on [x] is made salient, ⌜x *because* y⌝ is true, and that in a context in which intervening on [x] in order to intervene on [y] is made salient, a majority would judge that ⌜y *because* x⌝ is true. We found both predictions supported in all three conditions.

*Table 3. Results of* 𝜒*2-tests for level of truth agreement.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Condition** | **%Yes** | **%No/I** | 𝜒2 | ***p*-value** |
| **Mind/Brain** | | | | |
| **1A:**  **Stress because cortisol** (N = 74) | 71.6 | 28.4 | 13.838 | <.001 |
| **1B: Cortisol because stress** (N = 71) | 77.5 | 22.5 | 21.423 | <.001 |
| **Independence test:** |  | | .651 | .420 |
| **Sets** | | | | |
| **2A: Set because member** (N = 56) | 85.7 | 14.3 | 28.571 | <.001 |
| **2B: Member because set** (N = 64) | 71.9 | 28.1 | 12.250 | <.001 |
| **Independence test:** |  | | 3.370 | .066 |
| **Euthyphro** | | | | |
| **3A: Good because God approves** (N = 76) | 65.8 | 34.2 | 7.579 | .006 |
| **3B: God approves because good** (N = 80) | 67.5 | 32.5 | 9.800 | .002 |
| **Independence test:** |  | | .051 | .821 |

While the majority of people in all conditions thought that the relevant proposition was true, we thought that which kind of vignette participants saw might have an effect on the *relative* majority. In order to investigate this possibility, we performed separate 𝜒2 of independence tests for level of truth agreement within each kind of vignette. The results of these separate independence tests show that there were no significant differences in proportions between any pair of vignettes. That is, there is no difference in proportions of participants who agree that ⌜x *because* y⌝ or ⌜y *because* x⌝ is true between each kind of vignette in either the Mind/Brain case, Sets case, or Euthyphro case.

Next, we hypothesised that a majority of people would judge that at a condition in which intervening on [y] in order to intervene on [x] is made salient, ⌜x *because* y⌝ is an explanation for the relevant subject, and that at a context in which intervening on [x] in order to intervene on [y] is made salient, a majority of people would judge that ⌜y *because* x⌝ is an explanation for the relevant subject. We found both predictions supported in all three conditions.

*Table 4. Results of* 𝜒*2-tests for level of explanation agreement.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Condition** | **%Yes** | **%No/I** | 𝜒2 | ***p*-value** |
| **Mind/Brain** | | | | |
| **1A:**  **Stress because cortisol** | 71.6 | 28.4 | 13.838 | <.001 |
| **1B: Cortisol because stress** | 81.7 | 18.3 | 28.521 | <.001 |
| **Independence test:** |  | | 2.046 | .153 |
| **Sets** | | | | |
| **2A: Set because member** | 78.6 | 21.4 | 18.286 | <.001 |
| **2B: Member because set** | 75 | 25 | 16.000 | <.011 |
| **Independence test:** |  | | .213 | .644 |
| **Euthyphro** | | | | |
| **3A: Good because God approves** | 68.4 | 31.6 | 10.316 | .001 |
| **3B: God approves because good** | 73.8 | 26.3 | 18.050 | <.001 |
| **Independence test:** |  | | .539 | .463 |

While the majority of people in all conditions thought that the (relevant) proposition is an explanation for the subject, we once again thought that which kind of vignette people saw might have an effect on the relative majority. So in order to investigate this possibility, we again performed separate 𝜒2 of independence tests for level of explanation within each kind of vignette. The results of these separate independence tests show that there were no significant differences in proportions of people who judge that ⌜x *because* y⌝ or ⌜y *because* x⌝ is an explanation for the relevant subject between any pair of vignettes.

Finally, we hypothesised that there would be no significant difference between participants’ mean levels of agreement regarding whether ⌜x *because* y⌝ is true, and their mean levels of agreement regarding whether ⌜x *because* y⌝ is an explanation for the relevant subject. To test this we compared within-subjects’ level of truth agreement and level of explanation agreement with separate paired-samples t-tests, reported in Table 5 below. We found no significant difference between an individual’s level of truth agreement and level of explanation agreement across any of the 3 pairs of vignettes.

*Table 5. Results of within-subjects t-test for level of truth agreement vs. level of explanation agreement.*

|  |  |  |
| --- | --- | --- |
| **Level of Truth Agreement vs. Level of Explanation Agreement** | ***t-value*** | ***p*-value** |
| **Mind/Brain** | | |
| **1A:**  **Stress because cortisol** | .656 | .514 |
| **1B: Cortisol because stress** | 1.341 | .184 |
| **Sets** | | |
| **2A: Set because member** | 1.117 | .269 |
| **2B: Member because set** | 1.050 | .298 |
| **Euthyphro** | | |
| **3A: Good because God approves** | .405 | .686 |
| **3B: God approves because good** | -.076 | .940 |

**5 Discussion**

All nine of our hypotheses were vindicated. In all three cases a majority of participants judge that ⌜x *because* y⌝ is true, and is an explanation, in a condition in which intervening on [y] in order to intervene on [x] is made salient, and a majority judge that ⌜y *because* x⌝ is true, and is an explanation, in a condition in which intervening on [x] in order to intervene on [y] is made salient. This is also reflected in mean levels of truth and explanation agreement.

In some cases, these results were more surprising than others. Some of these judgements align with philosophers’ judgements about what explains what. The more surprising judgements, however, were in 3(a) that eating snails is good because God approves of it; in 1(b) that Fred has high cortisol levels because he is stressed, and in 2(b) that Eminem exists because the singleton set containing Eminem exists, all of which were judged to be true, and an explanation. Cases 1(b) and 2(b) are particularly noteworthy, since in each a majority of people make a judgement that is firmly at odds with philosophical judgements on the matter. Moreover, while some philosophers judge 3(a) to be the correct direction of explanation, as far as we know all philosophers think that *either* things are good because God approves of them, or that God approves of them because they are good, but not both. However, a majority of our participants make *both* judgements.

Notably, in all three cases a majority of participants make symmetrical judgements, contrary to philosophical orthodoxy about these cases.

Also of note was that there was no significant difference between participants’ mean levels of truth agreement regarding whether what Maria said in vignette 1(a), 2(a) and (2c) was true, and their means levels of truth agreement regarding whether what Maria said in vignette 1(b), 2(b) and 3(b) was true. That is, when we look at vignette 1(a) and 1(b), and look at participants’ mean levels of truth agreement, we find no significant difference, and *mutatis mutandis* for 2(a) and 2(b) and for 3(a) and 3(b). So, for instance, a majority of participants judged that ‘Eminem exists *because* the singleton set containing Eminem does’, is true, and is an explanation for Maria. Further, a majority judged that ‘the singleton set containing Eminem exists *because* Eminem does’, is true, and is an explanation for Maria. Even so, one would have expected participants to *more strongly* agree to the latter statement than to the former, even given the change of condition. Yet we found that people *as strongly* agree that the former statement is true as they do the latter. (*Mutatis mutandis* for the other two cases).

We find the same pattern of results when we look at participants’ mean levels of explanation agreement across the three pairs of vignettes. In each case, mean levels of explanation agreement between vignette 1(a) and 1(b) are not significantly different, and likewise for the remaining two pairs of vignettes.

These results show that metaphysical explanation, as non-philosophers employ the notion, is decidedly not asymmetric. Moreover, we don’t need to appeal to the sorts of recherché cases to which philosophers have pointed, to find examples of symmetrical judgements. A majority of participants make symmetrical judgements in every case we tested. This suggests that the breadth of these symmetrical instances is far greater than we might otherwise have thought. Indeed, our evidence is consistent with this ordinary notion of metaphysical explanation being symmetric rather than non-symmetric, though of course the number of cases tested would need to be far greater to support this conclusion.

Interestingly, though, our results leave open that people will judge metaphysical explanations to be *quasi*-asymmetric: that is, they leave open that when people judge that ⌜x *because* y⌝ is true, and is an explanation, in some condition, that they may well also judge that *in that condition* ⌜y *because* x⌝ is not true, and is not an explanation.[[11]](#footnote-11) So it may be that in this narrow sense people are not prepared to countenance circular explanations. This, we think, is something that could profitably be investigated in further research.

What ought we conclude about metaphysical explanation on the basis of this research? That depends on how one sees the connection between ordinary judgements about what metaphysically explains what, and our theorising about metaphysical explanation. One could argue that the fact that these judgements are so dissimilar to those of philosophers suggests either that participants are very confused, or that they are employing a quite different notion of metaphysical explanation to that of philosophers.

We see no real reason to think that participants are confused. Although some of the vignettes describe unusual circumstances, they are not difficult to understand. Perhaps, though, this is reason to think that the notion of metaphysical explanation employed by non-philosophers is quite different from that employed by philosophers. Our results have no implications for philosophers who take themselves to be developing a specialised notion that can be used to, *inter alia,* make sense of relations of relative fundamentality (say) and which is not taken to (or at least, which need not) connect to an ordinary notion of metaphysical explanation. Philosophers can define up a new notion however they see fit, and they need not be constrained in so doing, by the ways in which non-philosophers use a similar, or related, notion. Call this the *specialised notion approach*.

As we noted at the beginning of this paper, however, many philosophers take themselves to be employing an everyday notion and to be tidying up that notion, formalising it and then exploring it. For these philosophers our results have clear implications. They suggest one of two options. Option one is for these philosophers to re-evaluate the task they took themselves to be engaging with, and to see that they are in fact employing and developing a quite different notion of metaphysical explanation than that employed by non-philosophers. These philosophers would then come to embrace the specialised notion approach.

The second option is for these philosophers to maintain that the notion that are employing is, at heart, the same as that of non-philosophers, and to use empirical data such as that described here to inform their account of metaphysical explanation. Call this to take the *everyday notion approach*. To pursue this approach would require further empirical research into the conditions under which people judge that one thing metaphysically explains another, and then a way of systematising all those judgements to generate an account of the notion in question. Our results strongly suggest that any systematisation of our judgements will not yield an account of metaphysical explanation as asymmetric. Indeed, it suggests that the current ways of thinking about metaphysical explanation have over emphasised the role of asymmetry in our judgements. That, in turn, has significant down-stream consequences for theorising about the conditions under which propositions of the form ⌜x *because* y⌝ are true.

There are already a great number of these accounts, ranging from those that appeal to relations of ground in some way[[12]](#footnote-12), through to those that do not.[[13]](#footnote-13) Many, but not all, of the accounts that appeal to ground posit a primitive, asymmetric, relation, in part motivated by the purported need for some asymmetric worldly structure that can back asymmetric metaphysical explanation. But if metaphysical explanation is not asymmetric, then positing a relation with these features, (at least for these reasons) is clearly a mistake. Similar considerations apply to non-grounding approaches that attempt to capture the asymmetry of metaphysical explanation.

So if philosophers embrace the everyday notion approach, our results suggest that some of the theorising about the properties of the worldly structures that back metaphysical explanations are mistaken: for they are premised on erroneous assumptions about the properties of metaphysical explanations.

Notably, either approach requires some change in methodology or in argumentative rhetoric. The specialised notion approach leaves much of current philosophical methodology untouched; but it requires we concede that the notion of metaphysical explanation, as it is used by philosophers, does not have some long-standing pedigree: it is not a well understood, broadly used notion on which we all have a good grasp. Indeed, it leaves open the charge that philosophers have invented some new, parochial notion, which may be uninteresting, or worse, unintelligible.

By contrast, the everyday notion approach requires significant methodological upheaval. If philosophical judgements are not a good guide to everyday judgements, (as this research suggest) then much more engagement with empirical work is required. Moreover, there are then questions about why philosophical judgements are so different from everyday judgments if they are both judgments about the very same phenomenon, and in turn, questions about how best to systematise these quite different judgements into a cohesive account of metaphysical explanation.

Whichever approach is pursued, we take it that there is still interesting and important work to be done in investigating the everyday notion of metaphysical explanation employed by non-philosophers, whether this is taken to be the same, or different, from the notion employed by philosophers. We hope to have begun making some, small, and inroads into this task.

**References**

Audi, P. (2012a). ‘A clarification and defense of the notion of grounding.’ In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 101–121). Cambridge: Cambridge University Press.

Audi, P. (2012b). ‘Grounding: Toward a Theory of the In-Virtue-Of Relation’. *Journal of Philosophy* 109 (12):685-711.

Baron, S., & Norton, J.(2019). ‘Metaphysical Explanation: The Kitcher Picture.’ *Erkenntnis*. DOI 10.1007/s10670-018-00101-2

Bertrand, M. (2019). “Metaphysical Explanation by Constraint”. *Erkenntnis* 84 (6):1325-1340.

Bliss, R. (2014). “Viciousness and Circle of Ground”. *Metaphilosophy* 45 (2):245-256.

Bliss, R. (2013). “Viciousness’ and the Structure of Reality”. *Philosophical Studies* 166 (2):399-418.

Cameron, R. (2008). “Turtles all the way down: Regress, Priority and Fundamentality”. *Philosophical Quarterly* 58 (230):1-14.

Correia, F. (2014). “Logical Grounds” *The review of Symbolic Logic,* 7/1,pp. 31–59.

Dasgupta, S. (2017). ‘Constitutive Explanation.’ *Philosophical Issues* 27(1):74-97.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*(1).

Glazier, M. (2020). Explanation. In Michael J. Raven (ed.), *Routledge Handbook of Metaphysical Grounding*. London: pp. 121-132.

Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., and Danks, D. (2004). ‘A theory of causal learning in children: Causal maps and Bayes nets.’ *Psychological Review,* 111:1–30.

Hilton, D. J. (1990). ‘Conversational processes and causal explanation.’ *Psychological Bulletin*, *107*(1), 65.

Hilton, D. J., & Erb, H. P. (1996). ‘Mental models and causal explanation: Judgements of probable cause and explanatory relevance.’ *Thinking & Reasoning*, 2(4), 273–308.

Hitchcock, C., & Knobe, J. (2009). ‘Cause and norm.’ *Journal of Philosophy*, 11(11), 587–612.

Hoyos, C., & D Gentner (2017). ‘Generating explanations via analogical reasoning’ *Psychonometric Bulletin Review* 24:1364–1374 DOI 10.3758/s13423-017-1289-5

Jenkins, C. 2011. “Is Metaphysical Dependence Irreflexive?”, The Monist 94: 267–76.

Kahneman, D., & Miller, D. T. (1986). ‘Norm theory: Comparing reality to its alternatives.’ *Psychological Review,* 93(2), 136–153.

Kovacs, D. M. (forthcoming). ‘Metaphysically Explanatory Unification.’ *Philosophical Studies*. DOI:10.1007/s11098-019-01279-z

Latham, A. J., & Miller, K. (ms) “Judgements of Metaphysical Explanations are Context Sensitive” <https://philpapers.org/rec/LATJOM>

Lombrozo, T. (2007). ‘Simplicity and probability in causal explanation.’ *Cognitive psychology*, *55*(3), 232-257.

McGill, A. L. (1989). ‘Context effects in judgments of causation.’ *Journal of Personality and Social Psychology*, *57*(2), 189.

Miller, K, & Norton, J. (2017). ‘Grounding: it’s (probably) all in the head.’ *Philosophical Studies*, 174(12), 3059-3081.

Murphy, G. L., & Medin, D.L. (1985) The role of theories in conceptual coherence. Psychol. Rev. 92, 289–316.

Norton, J., & Miller, K. (2017). ‘A Psychologistic Theory of Metaphysical Explanation’. *Synthese*. DOI 10.1007/s11229-017-1566-x

Raven, M. J. (2012). “In Defence of Ground.” *Australasian Journal of Philosophy* 90 (4):687 - 701.

Raven, M. J. (2013). ‘Is ground a strict partial order?’ *American Philosophical Quarterly*, *50*(2), 193-201.

Raven, M. J. (2015). ‘Ground.’ *Philosophy Compass*, 10(5), 322–333.

Rodriguez-Pereyra, G. (2005). ‘Why Truthmakers?’ In H. Beebee and J. Dodd (eds.), *Truthmakers: The Contemporary Debate*. Oxford: Clarendon Press.

Rodriguez-Pereyra, G. (2015). ‘Grounding is not a Strict Order.’ *Journal of the American Philosophical Association,* 1(3), 517-534. doi:10.1017/apa.2014.22.

Rosen, G. (2010). ‘Metaphysical Dependence: Grounding and Reduction’. In *Modality: Metaphysics, Logic, and Epistemology*, edited by B. Hale and A. Hoffmann, pp. 109–36.

Schaffer, J. (2009). ‘On What Grounds What.’ In D. Manley, D. Chalmers, and R. Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Oxford University Press.

Schaffer, J. (2010). ‘The Internal Relatedness of All Things’ *Mind* 119:341-376.

Schaffer, J. (2016). ‘Grounding in the image of causation.’ *Philosophical Studies* 173 (1):49-100.

Schaffer, J. (2012.) “Grounding, Transitivity, and Contrastivity”, in F. Correia & B. Schnieder (eds.), Metaphysical Grounding: Understanding the Structure of Reality. Cambridge: Cambridge University Press.

Shaheen, J. L. (2017). ‘The causal metaphor account of metaphysical explanation.’ *Philosophical Studies*, *174*, 553–578.

Sloman, S. (2005). *Causal Models: how people think about the world and its alternatives*. Oxford: OUP.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive science*, *27*(3), 453-489.

Thompson, N. 2016. “Metaphysical Interdependence”, in M. Jago (ed.), *Reality Making.* Oxford: Oxford University Press.

Trogdon, K. (2013). An introduction to grounding. In Miguel Hoeltje, Benjamin Schnieder & Alex Steinberg (eds.), Varieties of Dependence. Munich, Germany: Philosophia Verlag. pp. 97-122.

Vasilyeva, N., Wilkenfeld, D., & Lombrozo, T. (2017). ‘Contextual utility affects the perceived quality of explanations.’ *Psychonomic bulletin & review*, *24*(5), 1436-1450.

Wilson, A. (2018). ‘Metaphysical causation.’ *Noûs*, *52*(4), 723-751.

Wilson, J. (2014). ‘No Work for a Theory of Grounding.’ *Inquiry*, 57(5–6):1–45.

1. We use corner quotes here to signify that ⌜x *because* y⌝ is a *kind* of sentence, where x and y are variables that range over sentences. We will speak of ‘an instance of ⌜x *because* y⌝’ when we intend to talk about a particular instance of the schema. We will simply speak of ⌜x *because* y⌝ in order to talk about all instances of the schema. [↑](#footnote-ref-1)
2. Though there are some, such as Wilson (2018) and Schaffer (2016), who think that metaphysical explanation is a sort of non-diachronic causation. [↑](#footnote-ref-2)
3. See Schaffer, 2009, 2010; Raven, 2013, 2015; Rosen, 2010; Audi, 2012a, 2012b; Cameron, 2008). [↑](#footnote-ref-3)
4. Schaffer (2012) denies that metaphysical explanation is transitive, Thompson (2016) and Bliss (2013 and 2014) deny it is asymmetric, Correia (2014) and Jenkins (2011) deny that it is irreflexive, and Rodriguez-Pereya (2015) denies it is any of the three. [↑](#footnote-ref-4)
5. Bliss (2013, 2014). [↑](#footnote-ref-5)
6. Sometimes these are framed as examples of symmetrical instances of grounding. Since authors often move freely between these terms, it is clear that these examples constitute examples of symmetrical metaphysical explanation. [↑](#footnote-ref-6)
7. The only work we know of is that by Latham and Miler (ms), which investigates whether judgments about metaphysical explanation are context sensitive. [↑](#footnote-ref-7)
8. Gopnik et al., (2004), Steyvers, Tenenbaum, Wagenmakers & Blum (2003) and Sloman (2005). [↑](#footnote-ref-8)
9. Einhorn & Hogarth (1986) and Murphy & Medin (1985). [↑](#footnote-ref-9)
10. These are people in a large database who partake in a range of online experiments, usually in psychology, behavioral economics and sociology, for monetary compensation. While they have significant experience in completing online experiments, there is little reason to think that these people will have a particular interest in, or knowledge of, philosophy. [↑](#footnote-ref-10)
11. Except in cases of reflexive instances of metaphysical explanation. [↑](#footnote-ref-11)
12. Schaffer (2009), Cameron (2008), Audi (2012a), Raven (2012) and Trogdon (2013) Fine (2012); Correia (2005). [↑](#footnote-ref-12)
13. For a variety of such approaches see Kovacs (forthcoming) Miller and Norton (2017) Baron and Norton (2019), Wilson (2014), Shaheen (2017), Bertrand (2019), Norton and Miller (2017). [↑](#footnote-ref-13)