HOW BELIEF WORKS

Derrick Farnell

Why do we believe what we believe? The answer may seem obvious: we believe what we've assessed to be true. However, there's a surprisingly basic logical flaw in this theory. And even more surprising is the real answer to the above question.

1. Why belief can't be the product of an assessment of truth

1.1 The problem

The content of any belief is a claim, whether it's something profound, like 'There's an afterlife', or something mundane, like 'There's milk in the fridge'. According to the above theory of belief-formation, we believe a particular claim – whether it was produced by our own mind or someone else's – because we've assessed that it's true. Indeed, the *Oxford English Dictionary* defines 'believe' as 'to consider to be true'.[1]

However, consider the following two claims:

There's milk in the fridge.

The claim 'There's milk in the fridge' is true.

They may seem to be saying the same thing. However, they aren't, because whereas the second claim refers to the first, the first obviously doesn't. That is, whereas the first simply refers to the existence of milk in the fridge, the second refers to a claim about the existence of milk in the fridge. They are instead merely 'logically equivalent' – that is, they merely directly imply each other. If there's milk

doi:10.1017/S1477175613000171
Think 35, Vol. 12 (Autumn 2013)

in the fridge, then the claim 'There's milk in the fridge' is true, and vice versa.

Therefore, although believing that there's milk in the fridge implies that we would conclude that the claim 'There's milk in the fridge' is true, and believing that the claim 'There's milk in the fridge' is true implies that we believe that there's milk in the fridge, they're not the same thing, contrary to the *OED*.

But it may still be thought that believing that there's milk in the fridge is a product of concluding that the claim 'There's milk in the fridge' is true.

However, first, if believing X was dependent on first believing that claim X is true, then believing that claim X is true would in turn be dependent on first believing that the claim 'claim X is true' is true, and so on, indefinitely. Belief-formation would therefore be impossible, and yet we obviously do form beliefs.

Also, by definition, the claim 'There's milk in the fridge' is true if, and only if, there's milk in the fridge. Therefore, in order to conclude that this claim is true, we must first believe that there's milk in the fridge – even if that belief was only formed immediately before reaching that conclusion. And if our conclusion that this claim is true is dependent on first believing that there's milk in the fridge, then the reverse can't also be the case.

That is, in order to conclude that claim X is true, we must first believe X, not the other way round.

1.2 Objection 1

It might be objected that there are at least some occasions when a belief is evidently the product of an assessment of truth. For example, if we believe that there's no milk in the fridge, but then someone says that there is, then this may lead to us assessing the truth of these contrary claims, which may then lead to us changing our belief.

However, in that scenario, what happens is this. In assessing the truth of these contrary claims, we may simply look inside the fridge. Upon seeing that there's milk

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

there, we form the belief that there's milk in the fridge – and the formation of this belief also constitutes the end of our belief that there's no milk in the fridge. After forming this belief, we can conclude that our original belief was false, and that the contrary claim, which we now believe, is true. Therefore, although the formation, or ending, of a belief may occur during an assessment of the truth of the claim in question, it nevertheless always occurs before the conclusion of that assessment, and so isn't a product of it. Instead, the conclusion of the assessment is a product of the belief or unbelief.

1.3 Objection 2

Another possible objection is that if we can only ever assess to be true what we already believe, then this implies that we can't form new beliefs, contrary to our evident ability to do so. However, the subconscious premise of this objection is that belief is the product of an assessment of truth. Given that beliefs, however they're formed, can't be the product of an assessment of truth, the fact that we can only ever assess to be true what we already believe doesn't itself imply that we can't form new beliefs.

1.4 Why we think that belief is the product of an assessment of truth

There are several possible reasons for the assumption that belief is the product of an assessment of truth. As explained:

1. We wrongly think that to believe X is, in itself, to consider claim X to be true.
2. Assessing whether a claim is true can stimulate the formation of our belief of it, but we don't notice that that belief actually formed before the conclusion of that assessment.

Also:

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

3. We forget that our beliefs are indeed merely beliefs, and think, when we're assessing the truth of claims, that we're comparing them directly with reality, and then forming our beliefs accordingly.
4. We know that if we conclude that claim X is true, then we believe X, and we then commit the logical error of confusing correlation for causation – that is, we conclude that the latter follows from the former causally, when it actually only follows logically, with the causal relationship actually being the reverse.

2. Another possibility

So why do we believe what we believe?

It might be thought that the answer has already been provided. In the scenario in part 1.2 we believed that there was milk in the fridge upon seeing milk in the fridge. So perhaps we believe X because we've perceived X.

However, a belief can of course also be the conclusion of reasoning. For example, we may believe that there's milk in the fridge at the present moment not because we can currently see that there's milk in the fridge, but because our memory of seeing milk in the fridge a few hours ago, and our knowledge that no one has been in the fridge since, leads us to this conclusion.

In fact, even beliefs that seem to be the direct product of perception are actually the conclusion of at least some reasoning. But that reasoning can be so basic, and therefore quick, that we don't notice it. For example, we never really see milk in the fridge, but merely a white liquid in a container. Our resulting belief that there's milk in the fridge is actually the conclusion of our reasoning about the nature of this white liquid. Indeed, we can believe contrary to what we perceive, such as when we know that an optical illusion that we're looking at is just that.

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

So it seems that we have the answer: we believe X because we've concluded X. In the case of a claim produced by someone else's mind, we'll therefore only believe it upon concluding it ourselves, subsequent to comprehending it. So, although belief isn't the product of an assessment of truth, it's the product of reasoning of some kind.

But even this is wrong. The real explanation of why we believe what we believe was first proposed by the seventeenth-century philosopher Baruch Spinoza[2] – although using a different argument from what follows – and has more recently received experimental support via the psychologist Dan Gilbert[3] and the neuroscientist, and writer, Sam Harris[4], among others. It's actually even simpler than any of the above theories, although much less obvious.

3. The origin of belief

3.1 Thinking X versus thinking about X

If claim X exists in our mind, then we must be either thinking X or thinking about X. For example, if the claim 'There's milk in the fridge' exists in our mind, then we're either simply thinking 'There's milk in the fridge' or we're thinking about the claim 'There's milk in the fridge'. In the first case we're simply thinking about the existence of milk in the fridge, whereas in the second we're thinking about a claim about the existence of milk in the fridge. The second case will consist of either the mental process of contemplating the claim 'There's milk in the fridge', or an outcome of that process – that is, a specific thought about this claim, such as 'He's wrong to think that there's milk in the fridge'.

Also, if we're simply thinking 'There's milk in the fridge', then we believe that there's milk in the fridge, whereas thinking about this claim doesn't necessarily mean that we believe it. For example, whereas thinking that this claim is true means that we believe it, thinking that this claim was made by a particular person earlier in the day doesn't.

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

It might be objected that while believing X necessarily involves thinking X, thinking X doesn't necessarily involve believing X. If it did, then we'd be able to form a belief in any claim just by choosing to think it, which obviously isn't true. For example, if we believe that there's no milk in the fridge, then we won't subsequently instead believe that there is milk in the fridge just because we've chosen to think, contrary to what we believe, the claim 'There's milk in the fridge'. However, thinking claim X indeed necessarily involves believing X.

3.2 To think X is to believe X

A claim is analogous to a painting of a scene, in the sense that both convey information about their subject, however accurately. That is, just as the content of a painting of a scene is a representation of that scene, the content of a claim is a representation of some aspect of reality. Indeed, we sometimes say that a particular description – a type of claim – 'paints a picture'. And the analogy applies even in this respect: just as a painting of a scene can, under the right circumstances, be confused for that scene itself, a claim can be confused for the aspect of reality that it refers to.

For example, imagine that a window frame is used to perfectly frame a painting of an outdoor scene, and the framed painting is then fixed to a wall inside a windowless room. If the overall effect is sufficiently realistic, then, upon entering the room for the first time, and seeing the outdoor scene in the painting, our initial visual experience of the painting may actually consist of seeing a real outdoor scene through a real window. In fact, if, as we're seeing the outdoor scene in the painting, our visual experience doesn't consist of seeing a mere representation of that scene, then the only other logical possibility is that it consists of seeing a real outdoor scene.

Regarding claims, as explained, if claim X exists in our mind, then we must be either thinking X or thinking about X. Therefore, as we're simply thinking 'There's milk in the

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

fridge', we're not thinking about this claim. That is, we're simply thinking about the existence of milk in the fridge, and not about this claim about the existence of milk in the fridge. Therefore, as we're simply thinking this claim, its content can't exist in our mind as the content of a claim, because that would involve thinking about the claim. And if, as we're simply thinking this claim, its content doesn't exist in our mind as the content of a mere claim – a mere representation – then the only other logical possibility is that it exists in our mind as reality. And to say that the content of a claim exists in our mind as reality is to say that we believe it. Therefore, simply thinking 'There's milk in the fridge' involves believing that there's milk in the fridge. And the same logic applies to our thinking any claim: thinking claim X involves believing X.

Also, given that believing X involves claim X being stored in our memory as reality, our belief of X doesn't cease when we finish thinking X. However, in the earlier scenario of choosing to think, contrary to what we believe, the claim 'There's milk in the fridge', within a fraction of a second of doing so we'll recall that we were thinking this claim simply because we'd chosen to think contrary to what we believed, and therefore not because we'd any reason to change our belief, and we'll therefore again think/believe 'There's no milk in the fridge'. And because our belief in the existence of milk in the fridge was so brief, we'll likely have no recollection of it.

The fact that thinking X involves believing X explains the attraction of self-improvement mantras. For example, if someone with normally low self-esteem repeatedly says, while also thinking, 'I'm a lovable person', and they can maintain enough focus on performing this task so that they're prevented from thinking anything else, then they'll believe this for as long as they do so, thereby giving them, albeit temporarily, the increased self-esteem that they crave. The effect is always temporary, because this practice doesn't address whatever is causing their low self-esteem. However, the temporary boost in self-esteem often leads to

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

the misplaced hope that regular sessions of this practice will eventually lead to a long-term increase in self-esteem, just as repeated exercise of a muscle increases its strength.

However, the fact that thinking claim X involves believing X doesn't completely answer the question of why we believe what we believe, because we still need to understand why we end-up thinking X instead of merely thinking about X.

3.3 Why we believe what we believe

Consider how a claim enters our mind.

A claim, X, entering our mind was either produced by our own mind or was produced by someone else's mind and then comprehended by us. Therefore, either way, X enters our mind as the output of a process in our mind.

This process, either way, isn't that of thinking about X, because the process by which X enters our mind involves, by definition, X existing in our mind only at the very end, whereas the process of thinking about X involves, by definition, X existing in our mind for its full duration. In the case of a claim that was produced by someone else's mind, it might be thought that the process of comprehending that claim involves thinking about it, but it actually involves an analysis of whatever is communicating the claim – speech sounds, text, etc. – in order to reproduce the claim.

So the process of thinking about X is dependent on X entering our mind as the output of a previous process in our mind. However, we can't be thinking about X at the moment that it enters our mind.

Again, if claim X exists in our mind, then we must be either thinking X or thinking about X. Therefore, the output of the process by which X enters our mind must be either the thought 'X' or the process of thinking about X. However, the content of the output of the process which produces, or reproduces, claim X is simply 'X'. Therefore, the output of the process by which X enters our mind must simply be the thought 'X', with the content of that output

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

then becoming the input to the process of thinking about X. And to think X is to believe X.

In short, we believe every claim that enters our mind, whether it was produced by our own mind or someone else's.

4. Why we don't seem to be completely credulous

You may find this conclusion hard to accept, even if you can't find fault in the reasoning. We certainly don't seem to always jump to conclusions when we're reasoning about something, and nor to be always gullible when we're comprehending claims produced by others. And we can think of countless specific claims, whether produced by our own mind or other people's, that we don't believe, despite them having entered our mind. Indeed, it would be logically impossible to believe every claim that enters our mind, given that many such claims contradict those that have previously entered our mind.

However, regarding our not believing specific claims that have entered our mind, the above conclusion isn't that we indefinitely believe every claim that enters our mind. It's simply that we believe claims upon them entering our mind, which doesn't exclude the possibility of us subsequently ceasing to believe them. It may be objected that there are many claims that we didn't believe even upon them entering our mind. But, as explained next, our capacity to unbelieve also explains why we don't seem to always believe claims upon them entering our mind, even though we do.

In the scenario in part three of choosing to think, contrary to what we believe, the claim 'There's milk in the fridge', our decision to think contrary to what we believe is in our mind immediately after thinking/believing this claim, which leads us to think/believe 'There's no milk in the fridge' immediately afterwards. And as explained, it's therefore likely that we subsequently won't seem to have believed the claim 'There's milk in the fridge' as we thought it, given

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

the briefness of our belief. Of course, normally, when a claim enters our mind – whether it was produced by our own mind or someone else's – and we're therefore thinking/believing the claim as it does, such a thought/belief isn't the product of such a decision. Nevertheless, such a belief may still be subsequently replaced by another – and possibly by an immediately preceding belief – so soon after its formation that we have no recollection of it.

Part one presented the scenario of believing that there's no milk in the fridge, but then hearing someone say that there is, which then led to our assessment of the truth of these contrary claims, during the course of which we changed our belief. What wasn't explained is that we'll actually believe this contrary claim upon comprehending it. However, we'll then unbelieve it if we recall our certainty, before hearing the claim, about the non-existence of milk in the fridge, and then therefore think/believe both 'There's no milk in the fridge' and 'They're mistaken'. And although it takes several seconds to read the previous sentence, the process described may only take a fraction of a second. Therefore, we may have no recollection of our brief belief in the existence of milk in the fridge. However, if we then consider the other person's own apparent certainty in the existence of milk in the fridge, then this may lead, within a further fraction of a second, to us having at least some doubt in the non-existence of milk in the fridge. That is, our quickly regained belief in the non-existence of milk in the fridge may also only last a fraction of a second, before being replaced when we think/believe 'It's uncertain whether there's milk in the fridge'. And that uncertainty may then motivate an assessment of the truth of these contrary claims.

Even if, before hearing this claim, we didn't hold a belief about the existence of milk in the fridge, our belief of the other person's claim, upon comprehending it, could still be very short-lived. For example, we could subsequently think/believe 'They could be wrong' – perhaps because we know that they haven't actually checked the fridge since

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

other people have accessed it – and we may then think/believe 'It's uncertain whether there's milk in the fridge'. And, again, such a thought process could take place within a fraction of a second of the formation of our belief in the existence of milk in the fridge, and we may therefore have no recollection of it.

To take a more extreme example, when we read the following:

> You're a tree.

we must believe this claim as we do. But our belief only lasts the fraction of a second that it takes us to conclude that we're not a tree, and we'll therefore likely have no recollection of it.

So, although we always believe a claim upon it entering our mind – whether it was produced by our own mind or someone else's – that belief can also then be replaced with equal ease, and possibly by an immediately preceding belief, and possibly within such a short period of time that we have no recollection of our brief belief. Therefore, upon being presented with this theory of belief formation, and then thinking about how we form beliefs in practice, we may falsely recall, or imagine, cases of us not believing claims upon them entering our mind. Also, as will be explained in part seven, the briefness of such beliefs is one of several reasons why our belief of every claim that enters our mind doesn't naturally come to our attention.

5. We don't believe X because we've concluded X

As with the theory that we believe X because we've assessed claim X to be true, the theory that we believe X because we've concluded X is actually the wrong way around: we conclude X because we're thinking/believing X as claim X enters our mind when we're reasoning. That is,

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

although claim X may be the product of our reasoning, our believing X isn't.

To be clear, this is not to say that the course of our reasoning is irrelevant to what we believe. The content of claims that enter our mind as the product of reasoning, and which we then believe, is obviously determined by the course of that reasoning, even if our believing those claims isn't due to them being the product of reasoning.

In the case of comprehending a claim, X, produced by someone else, X doesn't even enter our mind as the product of reasoning on our part, and yet we must believe it upon it entering our mind.

And even a claim entering our mind that was produced by our mind isn't necessarily the product of our reasoning. For example, as we're trying to work-out how there could be milk in the fridge, contrary to our belief prior to looking inside the fridge, an explanation, such as 'Someone put new milk in the fridge after I last checked', may enter our mind as the product not of our reasoning, but of our imagination, or as a memory that we'd previously failed to recall.

6. Belief is self-preserving

Another common misconception about belief – in addition to the theories that we believe X because we've assessed claim X to be true, or perceived X, or concluded X – is that beliefs can be held with different levels of confidence, not just certainty.

Consider again the claim:

> There is milk in the fridge.

To be confident, but not certain, about this claim is to believe:

> It is likely, but not certain, that there is milk in the fridge.

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

And this claim is incompatible with the first, given that the wording of the first excludes the uncertainty in the second. Therefore, believing the first claim must consist of certainty in that claim.

And the same logic can even be applied to the second claim. That is, although that claim concerns an uncertainty, the first 'is' expresses certainty about that uncertainty. Therefore, to be less than certain about the second claim would be to similarly hold a contrary belief.

In general, if we're less than certain about a claim, then the content of our belief is that claim with an uncertainty introduced into it, which is therefore contrary to the original claim. Therefore, belief consists of certainty in the claim in question. That is, we can have different levels of confidence in a claim, but we'll only believe it if our level of confidence is certainty. And if our level of confidence in a claim is below certainty, then we believe – are certain about – a contrary claim. Therefore, the only justification for announcing our certainty in the content of a belief, X, is to emphasise that we indeed believe X, as opposed to merely believing that X is a likelihood.

Belief is therefore self-preserving: once we believe X, we'll normally not be motivated to assess its truth, including checking for other possibilities, given our certainty. As explained in part one, although what we believe can't be affected by the outcome of an assessment of truth, the process of assessing the truth of a claim can stimulate the formation, or ending, of our belief in it.

Our certainty doesn't mean that a contrary belief, including simply a doubt – a belief in the uncertainty of the claim in question – is prevented from ever replacing the current belief. Such contrary claims – whether produced by our own mind or someone else's – can of course still enter our mind in the course of subsequent thinking, even if that thinking isn't aimed at assessing the truth of the original claim. And as the examples in part four showed, this can happen within a fraction of a second of the formation of the belief being replaced. However, equally, our likely lack of

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

motivation to assess the truth of a belief means that it may endure even if the formation of a contrary belief, including simply a doubt, would only require a very small amount of reasoning.

So, not only do we believe a claim upon it entering our mind, that belief consists of self-preserving certainty in that claim.

7. Why we often don't notice that we believed a claim upon it entering our mind

As stated in part four, the obvious objection to the conclusion that we believe every claim that enters our mind is that we don't seem to. That is, we don't seem to always jump to conclusions when we're reasoning about something, and nor to be always gullible when we're comprehending claims produced by others. And it was explained in part four that although beliefs form with such ease, they can also then be replaced with equal ease, and possibly by an immediately preceding belief, and possibly within such a short period of time that we have no recollection of our brief belief. Therefore, upon being presented with this theory, and then considering how we form beliefs in practice, we may falsely recall, or imagine, cases of us not believing claims upon them entering our mind. But the briefness of such beliefs is also one of several reasons why our belief of every claim that enters our mind doesn't naturally come to our attention.

We're normally only motivated to analyse how we formed a particular belief if we subsequently unbelieve the claim in question, and then conclude that that belief was false, and we want to learn from our mistake. Therefore, it's normally only then that we may realise that we jumped to conclusions as we were reasoning about something, or were gullible with respect to a claim produced by someone else. But, for many of our beliefs, we won't subsequently unbelieve the claim in question, because:

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

1. A significant proportion our beliefs must be either true or so close to the truth that we don't notice that they're not true, and so are likely to endure.

   First, just because we believe a claim upon it merely entering our mind doesn't mean that that belief will be false. For example, if we see someone who looks exactly like someone we know, and the first possibility that enters our mind, which we therefore believe, is that this person is that person that we know, then this belief will usually be correct. That is, while we do sometimes misidentify people, such mistakes are uncommon relative to the total number of people that we initially correctly identify each day. Likewise, with respect to comprehending claims produced by others, when we, for example, believe someone's remark that they had an enjoyable meal the previous evening, in most such cases this will simply be a truthful claim.

   And our daily lives involve a constant stream of belief-formation. Consider only perceptual beliefs. We must constantly form beliefs about the world around us as we receive a constant stream of perceptual information from it, and as we interact with it. For example, even as I'm simply working at my desk typing these words, I'm forming beliefs about the positions of my fingers on the computer keyboard, about what letters are appearing on the screen, about the source of a melodic whistling sound coming from outside (a bird), and so on. And the same is true of our social beliefs: as we interact with others, we're constantly forming beliefs about what people are trying to communicate; what they're thinking and feeling, and why; how they might react to hearing what we're thinking of

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

saying next; and so on. Therefore, given our general ability to navigate the world around us, physically and socially, on the basis of our beliefs – albeit to varying degrees of success – a significant proportion of those beliefs must be true, or at least relatively close to the truth.

2. We sometimes never unbelieve a false claim.

   Atheists and theists will at least agree on this, and it's also true of more mundane beliefs. For example, if someone does lie about having had an enjoyable meal the previous evening, then we may never learn that this claim, which we believed, was untrue. Indeed, as with this example, many of the beliefs that we form concern things that we may never think about again.

   Also, as explained in part six, given that belief consists of certainty about the claim in question, and we're therefore normally not motivated to assess its truth, beliefs are self-preserving. And they can also be self-preserving in another sense, and simultaneously self-reinforcing. For example, if we falsely believe that someone has a selfish personality, but then observe them performing an act which is apparently altruistic, we may falsely conclude that they're actually performing the act for selfish reasons – such as to benefit from the resulting gratitude. So, our prior false belief led to an interpretation of this observation that not only protected the belief from being undermined by this observation, but also reinforced it.

Also, even for those claims that we do unbelieve, we may not become aware that we believed the claim in question upon it entering our mind, because:

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

1. As explained in part four, the existence of such beliefs may be so brief that we subsequently don't recall their existence.

   Therefore, we won't analyse their origins, especially if there wasn't sufficient time for them to form the basis of further false beliefs.
2. Even if we do subsequently recall a false belief, we're not actually guaranteed to analyse how it formed – and the more mundane the belief, or time-pressured we are, the less likely this is.
3. Even if we do attempt to establish how a false belief formed, we may be unable to, because we've forgotten, especially if it's some time after the formation of that belief.

Therefore, for at least these five reasons, the fact that we believe every claim that enters our mind doesn't naturally come to our attention.

8. Another objection

It might also be objected that when claim X first enters our mind, whether it's produced by our own mind or someone else's, what actually enters our mind isn't simply X, but something like 'It's possible that X' or 'They're claiming that X'. Therefore, even if we do believe a claim upon it entering our mind, what we believe when claim X first enters our mind isn't X, but merely these claims about X – that is, that X is a possibility, or that X is a claim that's been made by someone. Therefore, believing X is still dependent on producing X via some reasoning after it has entered our mind.

However, such claims about claim X – that X is a possibility, or that X is a claim that's been made by someone – are the product of an analysis of X. That is, these specific thoughts about X can only enter our mind as the

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

conclusion of the mental process of thinking about X. And as explained in part three, we can't think about X without first thinking/believing simply X. Therefore, claims never enter our mind as merely the subject of another claim, but always as standalone claims.

It might then be objected that in the case of comprehending a claim that's been produced by someone else, if this claim, X, is about another claim, Y, then comprehending X merely involves believing this claim about Y, and therefore doesn't involve believing Y. That is, while the person who produced this claim about Y must have at least begun that process believing Y, however briefly, what enters our mind is merely the output of that process. Indeed, if X and Y are contradictory – such as 'He's wrong to think that there's milk in the fridge', with Y being 'There's milk in the fridge' – then we can't simultaneously believe X and Y upon comprehending X. Therefore, claims can at least enter our mind as merely the subject of another claim, rather than as a standalone claim, when such a claim about another claim was produced by someone else. And we'll therefore not believe the claim that's the subject of the other claim, even briefly, as it enters our mind.

However, given that claim Y is the subject of claim X, we must first comprehend Y in order to comprehend X. That is, in order to comprehend the claim 'He's wrong to think that there's milk in the fridge', we must first comprehend the sub-claim 'There's milk in the fridge'. And as we comprehend the latter we'll indeed believe it, but we'll then unbelieve it when, within a fraction of a second, we combine our comprehension of it with that of the claim about it. Therefore, even when we're comprehending a claim about another claim that was produced by someone else, the latter claim must enter our mind as a standalone claim.

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

9. The origin of the availability bias

To recap:

- We believe every claim that enters our mind, whether it was produced by our own mind or someone else's.
- And given that belief consists of certainty, beliefs are self-preserving: once formed, we'll normally not be motivated to assess their truth, including checking for other possibilities.
- Our certainty doesn't mean that a contrary belief, including simply a doubt, is prevented from ever replacing the current belief. Such contrary claims – whether produced by our own mind or someone else's – can of course still enter our mind in the course of subsequent thinking, even if that thinking isn't aimed at assessing the truth of the original claim. And this can happen within a fraction of a second of the formation of the belief being replaced.
- However, equally, our likely lack of motivation to assess the truth of a belief means that it may endure even if the formation of a contrary belief, including simply a doubt, would only require a very small amount of reasoning.

All of this explains a phenomenon that's well-known to psychologists: the availability bias. This is our well-established disposition to form beliefs on the basis of whatever happens to be immediately available to our mind – that is, to jump to conclusions when we're reasoning about something, and to be gullible when we're comprehending claims produced by others, instead of suspending belief until we've checked for other possibilities. As the psychologist Stuart Sutherland wrote in his book *Irrationality* (using the term 'availability error' instead):

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

> . . . the inability to suspend judgement is one of the most prevalent aspects of irrationality.[5]
>
> . . . Judging by the first thing that comes to mind is called the 'availability error'. I have made it the first error described because it *permeates all reasoning* [my emphasis] and, as we shall see throughout the rest of the book, many other specific errors are in reality just further instances of it.[6]

The currently generally-accepted explanation of the availability bias is that it's due to a specific psychological mechanism which evolved to prevent us from becoming mentally paralysed by the cognitive demands of forming the constant stream of beliefs mentioned in part seven. However, the conclusion here is that it's due simply to the way that we – and any intelligent being – must form beliefs.

Regarding claims produced by our own mind, consider the everyday task of interpreting the tone of emails that we receive. Given the absence of both voice intonation and facial expression from this form of communication, the intended tone of an email is always open to at least some interpretation – a problem that tends to increase as our familiarity with the sender decreases. For example, is the tone of a reply consisting solely of 'Thanks' sincere or sarcastic? We'll believe whichever interpretation we think of first, and will then normally not be motivated to consider other possible interpretations – although another possibility, or simply a doubt, may still subsequently enter our mind. Therefore, unless the initial interpretation is obviously false for some reason, our premature conclusion has a reasonable chance of enduring. Thus, we believe, and are likely to continue to believe, the interpretation that happens to be immediately available to our mind, instead of suspending belief until we've checked for other possibilities.

Indeed, in one study,[7] participants correctly interpreted the tone of emails only 56% of the time on average – not

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

much greater than chance – and yet they guessed, on average, that they'd achieved a 90% success rate.

And we're just as vulnerable to the availability bias while sending emails: as we're composing an email, its intended tone is obviously immediately available to our mind, unlike alternative tones which the recipient could misinterpret the email as having. Indeed, the emails in the above study were sent by other participants, who guessed, on average, that the recipients would correctly interpret the tone 78% of the time.

Regarding claims produced by others, consider the following study.[8] Participants were asked to perform a particular task, and were then given feedback on their performance. However, between completing the task and being given the feedback, they were clearly warned that the feedback wouldn't actually be determined by their performance, but would be completely arbitrary. Nevertheless, when, after they'd received the feedback, they were asked to estimate how well they'd done, there was a correlation between such estimates and the feedback that they'd received. Given that we always believe what we comprehend, the participants will have believed the feedback as they heard it – that is, they believed the feedback because it was immediately available to their mind – and were therefore not as influenced by the preceding – and therefore less available – warning as they should have been.

10. The presumption of truth

So, just as the law has the presumption of innocence – the principle that everyone is considered innocent until proven guilty – so we implicitly presume that every claim that enters our mind – whether it was produced by our own mind or someone else's – is true, in the sense that we automatically believe it – with all the profound implications that such self-preserving premature beliefs have for how we think and act.

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press

Derrick Farnell has studied physics, philosophy and psychology, at St. Andrews University, the University of Western Ontario, and Edinburgh University. He currently works for the latter's main library. His main interest is psychology, specifically in relation to reasoning, happiness, ethics and politics. This article is also available on his personal website: tryingtothink.org. Email: derrick.farnell@gmail.com.

Notes

[1] 'believe, v., 4a', *OED Online* (Oxford University Press, March 2013), accessed March 25, 2013.

[2] Baruch Spinoza, *Ethics* (1677), Part 2, 49th proposition.

[3] D. Gilbert, 'How Mental Systems Believe', *American Psychologist*, vol. 46 (1991), 107–119.

[4] S. Harris *et al*., 'Functional Neuroimaging of Belief, Disbelief, and Uncertainty', *Annals of Neurology*, vol. 63 (2008), 141–147.

[5] Stuart Sutherland, *Irrationality* (London: Pinter & Martin, 2007), 1.

[6] Ibid., 11–12.

[7] L. Winerman, 'E-mails and egos', *Monitor on Psychology*, vol. 37, no. 2 (2006), 16.

[8] D.M. Wegner *et al*., 'The Transparency of Denial: Briefing in the Debriefing Paradigm', *Journal of Personality and Social Psychology*, vol. 49, no. 2 (1985), 338–346.

https://doi.org/10.1017/S1477175613000171 Published online by Cambridge University Press