

## Zombie and Qualia debate (Philosophy Forum, Aug-2004)

Arguing for the actual possibility of zombies and dualism is AKG in red and against Paul Knierim in blue.

It is my position that the existence of zombies is possible. Before we go any further, I'm sure the readers will like to know what is meant by the term "zombie," as it is not being used in a conventional sense.

A being that behaves like us and may share our functional organization and even, perhaps, our neurophysiological makeup without conscious experiences or qualia.

In order to show the possibility of a zombie, I will attempt to provide a consistent model for the relationship between consciousness and the human body (along with human behaviour) that allows for the existence of the human body and its behaviour without consciousness, or experience of qualia. Essentially, I am equating logical consistency (as well as non-contradiction of scientific facts) with possibility.

First, we all can agree on the difference between redness and electrical signals in the brain. That entity which experiences redness (qualia) and not electrical signals in the brain is the consciousness. The experience of redness by the consciousness may be caused by the electrical signals in the brain and the eyes' detection of electromagnetic radiation of a certain wavelength, but the two are not one and the same. This distinction is important.

It is also key to notice that the perception of redness is not material like the electrical signals in the brain. The image of the computer and desk in front of you is not any type of matter or energy we know of, although the brain and the electrical signals in the brain are matter and energy that we know of. Therefore, the existence of the brain does not immediately necessitate the existence of the consciousness; they are of differing nature and it remains to be shown that consciousness necessarily follows from the brain.

Clearly, my argument is that it is possible that the consciousness is not necessarily implied by the brain. That a brain can be a functioning one without giving rise to some consciousness. Again, please note that we are regarding consciousness as not simply some sort of self-awareness, but something that experiences qualia, qualia themselves being something different than the matter and energy said to give rise to them. The consciousness is an independent entity, an observer, and the brain only creates qualia if it has an associated conscious entity to perceive them. That is that the consciousness is necessary for the brain to create qualia, but the consciousness is not something that necessarily follows from the existence of the brain. The brain is sufficiently complex for us to admit the possibility that consciousness has origins other than the brain.

Essentially, this is not far from a solipsistic or skeptic position, one that suggests that we can't be certain of the existence of other minds, and that we might be the only conscious entities surrounded by beings who appear to behave as we do, consciously, but are really just very "complex machines" of some sort – in fact, the correct term would be zombies. Of course, this is not necessarily the most appealing or useful position, but the solipsist's position is a possible one, and that is all that I expect one to accept.

Now, we still need to tackle the question of behaviour. Can humans behave as they do without consciousness? Is the brain alone capable of emotion, appreciation of beauty, apparent unpredictability, and all the other things we normally ascribe to conscious action? I feel it is safe to say that this is indeed possible at very least, if not the fact. More and more, scientific studies are showing how much of our behaviour can be reduced to mental processes. We can indeed propose a consistent hypothesis that the brain is an incredibly complex "machine" which takes all sorts of input: visual, speech, ideas, values, emotions, etc, and returns with a seemingly conscious response. Again, I feel safe in saying that it is possible to explain human behaviour through mental processes without the intervention of a conscious entity.

In concluding my opening argument, I would simply like to summarize that it is indeed possible that the brain doesn't necessarily cause consciousness to arise, and that human action and behaviour can all be said

to be results of the incredibly complex workings of the brain, and so it is indeed possible for zombies to exist.  
By AKG

Philosophical zombies are people who are exactly like us in every physical respect, from their behavior down to their exact brain patterns. No matter what tests you perform, be they getting into long personal conversations or running brain scans, you will never be able to tell if your best friend is actually a zombie who does not possess any consciousness.

A class of dualists (henceforth zombie dualists) tell us that these zombies are entirely possible. They argue that there is no way to rule out the possibility that someone is a zombie. They also conclude that if zombies are possible, dualism must be true – this is an implication which I do not debate, I only argue that the condition is not fulfilled. Why do zombies imply dualism, you may ask? Put simply, the zombie scenario requires that consciousness be something mysteriously undetectable, and that it furthermore be independent from the physical world in the sense that facts of the physical world do not have to translate to facts of consciousness. The core of the zombie thesis is that knowing all of the physical facts cannot tell you for certain that there are any corresponding mental facts of consciousness.

Most everyone (even the zombie dualists) will concede that in actuality there aren't really hordes of zombies walking among us. We must ask how it is that we know this, how we have come to make this bold assertion that zombies are at least fairly unlikely. We could say that we know this because where we find sufficient brain activity of a particular type we know that there's consciousness. The zombie-advocate, however, must argue against this and say that it's only usually the case that the brain activity means consciousness and that it can by no means be taken as a necessity or rule. The zombie-advocate concedes that there aren't actually any instances where the brain activity isn't associated with consciousness, but pushes on to argue that there could be such cases. It just so happens that they always seem to come together so far... what an amazing coincidence, and such a fortunate one! Yet, how can we possibly know that we have been lucky enough to have this coincidence occur if as the zombie dualist tells us physical investigation doesn't prove consciousness? Much as the zombie dualist protests, we must recognize that there is actually nothing coincidental about it – it's necessity, not coincidence. There's no mysterious causation at work here which can be broken by putting your hand over the emitter. Consciousness doesn't just mysteriously affix itself to brain patterns, but rather the brain patterns are a perceptual representation of the consciousness itself.

Perhaps my opponent will suggest the possibility of zombies based on an epistemological argument regarding the problem of other minds. The claim which he could make is that we have only one example (the self) of a correlation between physical structure (brain patterns) and mental experience. To show via induction that this must hold true for all cases, he might suppose, requires a lot more examples than one. There's a major flaw in this argument. We should observe that the number of possible counterexamples to our one example of the correlation of consciousness with brain patterns is zero. There is no meaning to claiming that something is not true if there is no possibility of a counterexample to it... it's the same as saying that there's an invisible intangible undetectable pink elephant standing on someone's head. So, if we take the zombie scenario as it's presented, it is in a way self-defeating because according to it we could never verify that anyone does not have conscious experience. The zombie-advocate argues that how a person acts and how their brain appears under a CAT scan or any other sort of investigation is not a valid criteria for determining if they are conscious... he must therefore hold that it is just as impossible to prove someone is a zombie, as to prove that someone isn't. The zombie hypothesis resembles a global skeptical hypothesis in this way, and many of the objections to global skeptical scenarios could work against it.

We can certainly imagine the possibility of zombies, but this does not imply that zombies are actually possible. Conceivability does not imply possibility.

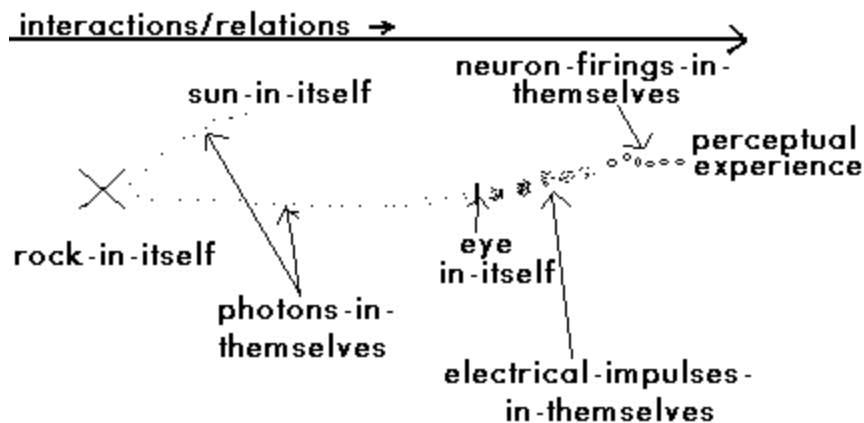
By Paul Knierim

I'll keep this one short and get straight to it. The first argument Paul presents is that consciousness is necessarily entailed by certain brain patterns. That it is not something of different origins which affixes itself to the brain, but something which is causally necessitated by brain patterns. There is, however, no real argument for this; as far as I can tell it was simply asserted. But could it even be argued for, could it be proven and scientifically tested? Could we scientifically test for consciousness? I think if not for certain, it is at least possible for it to be true that subjective experience of qualia cannot be tested, as qualia themselves are not matter or energy (the brain may not cause consciousness, but if there is a consciousness affixed to the brain, then the matter and energy in the brain may cause qualia, but aren't themselves qualia). If we can't directly test for consciousness, then we might want to see if consciousness is reflected in the behaviour, but as I've already argued, the behaviour could be entirely determined by the brain without experience of qualia.

Moreover, I believe Paul argues that the zombie dualist can fairly say that he only has one case where the mind and brain are related (the self) and no cases where it is known whether the brain is related to a consciousness or not (everybody else -- the problem of other minds). His argument here is that we have only one case for a brain related to a mind, but zero cases against the fact, and for some reason, he claims that this is an inductive argument in favour of proving the **impossibility** of zombies. One case is by no means sufficient for anything that could remotely be called an inductive argument. Moreover, even if we had 5 cases, Paul's goal is to prove the impossibility of zombies. If an inductive approach could ever be considered sufficient to prove with certainty the impossibility of something (which is not something I wish to argue), 1 case doesn't count as an inductive argument.  
By AKG

In his opening statements, AKG puts forth the classic qualia argument. Consciousness is what experiences the qualia of redness, rather than electrical signals in the brain, he tells us. Qualia is what it's like to be something. Thomas Nagel famously argued that the non-physical aspect of reality is the what-it's-like-to-be-iteness. What it's like to be a bat is not contained in any observations of the bat. No matter in how much detail we pin down the physical structure of the bat, we will never know what it is really like to be the bat. Nagel correctly observes that the only way we know what it's like to actually be things is by analogy with the self -- he chose a bat for his example because bats are enough like us that we're confident that there is *something* it's like to be one, but have different sensory organs so that we really cannot imagine *what* it would be like.

AKG's arguments seem to rest on the idea that it is somehow unique and surprising that the qualia of redness is different from electrical signals in the brain. To find that being something is different from observing something should not be surprising or unexpected. Electrical signals are observations, mappings of data collected by machines which detect the effects which the interactions of particles from your brain have on their instruments -- and then read by human eyes, which use a similar process themselves. It is absolutely impossible for anything more than relations to be preserved in these transfers of information. A quick illustration I made for a post a couple of years ago:



Each step along the way has the information being conveyed by a very different sort of carrier. Note that consciousness is neuron-firings-in-themselves (*not* neurons as they appear in our physical representations), this is all we actually experience. Clearly the state of the information about the rock -- or about the human brain being observed, if you wish to put that in place of the rock -- in consciousness will contain no more than a set of relationships in common with the original rock. It cannot possibly contain the intrinsic nature of the thing being observed, it cannot possibly contain anything beyond what can be conveyed by sets of relations.

*Quite coincidentally*, science has firmly established that neutron firings have a set of relationships in common with our conscious experiences – our qualia. Poke a person's brain with an electrical stimulator in a certain way, a certain place, and their qualia in their consciousness is altered in a corresponding way. AKG has proposed that this a coincidence, but examined objectively is not the suggestion that this correlation is coincidental exactly as absurd as to suggest that all the information I examine on my computer screen does not imply that my computer is turned on? I ask AKG if he will admit that there is no difference between the two claims, and if he will not then I ask him to clarify the difference.

The difference between the qualia of redness and the corresponding neuron firings simply establishes that observations are fundamentally different from inner experiences. In Kantian terms, the form of outer intuition differs from the form of inner intuition. All AKG has established by this line of argument is that you are not the world -- you are only yourself, containing a representation of the world. The thing in itself is not contained within the thing as it appears. To make zombies possible, the thing in itself would have to be *not necessarily related* to the thing as it appears... but surely we can see it would be nonsense to suggest that things in themselves and their appearances are unrelated.

I must note that I disagree with AKG's suggestion that the brain creates qualia. Qualia is what it's like to be the brain, it is not created by the brain. Things do not 'create' what it's like to be themselves.

AKG argues that electrical signals in the brain and conscious experience are not one and the same, and that the distinction is important. I fully agree. It is of fundamental importance to realize that the electrical signals, being as they are derived from observations, cannot be anything more than derivations of relative perceptual data. We do not jump outside of ourselves and grab the world-in-itself and then experience it as it is... the physical world is inherently grounded in the limitations of perception, with the most fundamental limitation being the subject-object distinction. When we make a physical observation, we necessarily create the subject-object dichotomy. The side effect of making an observation is that there is necessarily an observer involved (as becomes most unavoidably obvious in quantum mechanics, but is true everywhere).

Moving onward now to AKG's reply to my own post, I must first note that he either misstates or misleadingly states my position. I certainly do not argue that brain patterns cause consciousness, as he seems to be saying I do. Brain patterns are relative descriptions of consciousness. It would be absurd to say that a perceptually-derived relative description of something creates the actual thing -- a hard-core idealist

toeing the "to be is to be perceived" line might indeed argue that perception creates the world, but I certainly am not arguing for that.

My opponent noted that I had not yet given an argument yet for my position on the relationship of the mind and brain (neutral monism). I could and in fact already have written a book on this, but it will have to be compressed considerably for the purposes of this debate. I hope that this post has at least laid the groundwork of neutral monism, although thanks to the 1500 words per debate post limit I can't get into it very deeply here.

In reply to AKG's question in his second post, we can and most certainly do scientifically test for consciousness. A practical illustration of testing for consciousness consists of hitting a person over the head with a baseball bat. After doing this to a person, we say that they are **unconscious**. We say this for a reason -- not because we're deluded and have forgotten that we don't actually know when someone is conscious or not, but because we use our senses to test the victim for consciousness and conclude from the physical tests that there is none there. (Note that these unprofessional tests can be inaccurate. Sometimes the person is conscious but shows no visible signs. In *all* cases, however, sufficient scans of their brains with the right equipment do indeed yield a proper conclusion on if they're conscious or not. This technology already exists.)

As predicted, my opponent has also now argued in his second post that one case of consciousness is by no means sufficient to draw conclusions. (As we will see shortly, this again hearkens back to his eagerness to embrace global skepticism.) As already noted, there are zero possible counterexamples to the association of brain patterns with consciousness. AKG has not debated this, he seems to agree. So, we must conclude that he believes it makes sense to say "This [zombies] is possible, although it is absolutely theoretically impossible to ever gain evidence for it, and we already have a demonstration of counterevidence against it."

I would happily agree that there are things which can't be demonstrated which have counter-evidence but are still meaningful claims -- but those are all things which are, *in principle*, **possible** to demonstrate. Making an assertion which itself denies any possibility of evidence for the assertion is merely an exercise in sophistry.

Take a borderline example to help clarify:  
*Demon spirits inhabit all green objects.*

This would be meaningful if demon spirits can actually alter the object occasionally so that we could observe evidence of them, or even perhaps if we're assured that we can interview the demon spirits in the afterlife. If the demons are placid and are incapable of doing absolutely anything, however, it would be nonsensical to try to assert their existence. If your way of defining the demon spirits has just denied you all possible ways to know of them, you're being irrational.

It's a simple point: a rational person cannot assert something without having at least the *possibility* of evidence or a criteria for evidence on the matter. The zombie-advocate fails this test, the zombie-advocate is irrational.

The zombie dualist is the global skeptic in another guise, advocating a position which he admits it is absolutely theoretically impossible for him to ever have reason to believe. It is no better than Descartes' evil demon or Lehrer's Googols... the zombie as well is self-defeating, denying all possible evidence for itself.  
By Paul Knierim

Paul states that if we poke a person's brain with an electrical stimulator, we experience qualia accordingly. This is what I meant in saying that the brain causes qualia. Changes in the brain effect changes in perception. I don't recall claiming that this was coincidence. It seems to me that he has misinterpreted what I've said, so I will reiterate more clearly the mind-brain model I proposed.

Paul believes that I have misinterpreted his argument in saying that he claims that brain patterns cause consciousness. I will clarify here that there is distinction between subject and objects. The electrical signals sent to the brain are distinct from the brain. Similarly, the qualia-objects are distinct from the mind-subject (or perceiver or observer, whichever term you feel most comfortable with). Accepting the word of science, I will assume that if I am poked in the brain I will have a certain experience. That is, qualia necessarily follow from brain signals. But does consciousness, or the mind/perceiver follow **necessarily** from some aspect of the human physiology? If I dial a number 555-0101 ("555" numbers are fake), I that doesn't mean that it will reach a real person. The brain signals can do what they do, but if there is no conscious entity "affixed" to the brain then it won't necessarily generate qualia. In a sense, the existence of the consciousness is a contributing factor to the creation of qualia by brain patterns -- it is something of a causal agent allowing brain patterns to create qualia.

Paul continues claiming that we have a scientific test for consciousness. We use our senses to determine if the person is conscious (after we hit him over the head with a bat). However, just as the qualia are non-physical, so is the consciousness, and using senses to test something non-physical seems to be an unacceptable approach. Moreover, this argument implicitly assumes that human behaviour is a necessary indicator of consciousness. That is, if a person acts conscious, he's conscious. Of course, that's not what the zombie debate is about. The issue is that maybe some beings can act entirely like me, but not experience qualia. A very highly advanced A.I. may one day be able to replicate human behaviour, but will it experience qualia? Will it say process input (like red light) in a very complex manner and produce the output, "I see red," just like I would, but without experiencing the qualia of redness whereas I would? It seems that we can have various definitions for consciousness. We can detect something like a "physical consciousness" by scanning the brain, and this may tell us if a being will behave as we do, and it isn't hard to think that in the near future, machines will have this type of consciousness. But the relevant type of consciousness we're talking about is a non-physical one, the one that actually has the experience of qualia, not the one mirrored in the electrical patterns in the brain. This is something physical technology cannot really test for.

Paul also misinterprets possibility with necessity. Paul has the much harder position to defend; that zombies necessarily do not exist, i.e. all beings like humans have consciousness. We have no examples that we know of where humans exist without consciousness. I have one example of a human that does have consciousness: myself. However, this is not evidence against the possibility of zombies, it is evidence against the necessity of zombies in all cases of humans. Paul's argument is equivalent to saying that if we flip a coin once and note that it lands heads, we have some sort of evidence suggesting that it is impossible for it to land tails. We have evidence, in fact proof, that it is not necessarily going to land tails in all cases, but **by no means whatsoever** could we say we have evidence suggesting that tails are impossible.

Finally, Paul asserts that the zombie dualist is irrational because his claim is untestable. However, if we I claim that the existence of other minds is not testable, then whether you hold the position that no other minds exists or that other minds certainly exist, you are irrational. That is exactly the stance Paul is taking, one that claims certainty where certainty cannot be had. He claims that zombies are impossible, in part implying that other minds necessarily exist. All I claim is that it is not known, and this allows for the claim that it is **possible** for zombies to exist. I do not even claim that it is impossible to test for consciousness, but currently we have no means to do so.

In the same argument, Paul claims that if the demons have no effect on the green objects, it's ridiculous to assume they exist. This is precisely the point. If consciousness has no effect on human behaviour, how do we know it's there, and what reason do we have to claim that it is there? This is where the question of behaviour comes in: can a human being without consciousness (that which has the ability to experience qualia) act just like me? And again, I say that the brain is sufficiently complex that it cannot be claimed impossible that a human being will act just like me by virtue of the complexity of his brain, and not necessarily because he has a consciousness.

[A quick summary:](#)

- Human behaviour can be determined entirely by the brain, so human behaviour cannot be used as an indicator for consciousness, maybe only for a brain.
- Consciousness does not necessarily arise from the brain. I proposed a possible model which suggests essentially that the observer is an independent entity from the brain, and if, for whatever reason, an observer becomes "affixed" to the brain, the observer will experience qualia determined by the patterns in the brain.
- The question is if a being with the same physiology (especially brain and brain activity) and behaviour as myself can **possibly** exist without a consciousness, i.e. without experience of qualia. Since I've shown that it is **possible** for the behaviour to exist without the consciousness (as the brain may determine the behaviour and not the consciousness) and it is possible for the brain and brain activity to exist as it does without the consciousness (as the consciousness is possibly an independent entity).
- Since I don't intend to present a sophistic argument, I present not only a position that is logically consistent, i.e. conceivable, but one that it is indeed possible, and I do so by presenting one that does not contradict any scientific facts. The closest "fact" I have been presented with is that there exists at least one case of a human with a consciousness, that is myself. This is not a scientific fact, and no scientist would accept an experimental conclusion based on a sample size of 1. Moreover, I am not being difficult by limiting the sample size to 1, i.e. I don't claim that it is impossible to determine if other minds exist, but it is clear that hitting someone on the head or scanning their brains is not a sufficient test.

By AKG

AKG tells us in his summary that he "proposed a possible model" whereby consciousness does not necessarily correspond to all the physical indicators, but I must beg to differ. Declaring that you think something is possible does not constitute proposing a model. He has given us no model. He says himself in the same sentence where he calls it a model that this completely unexplained mysterious fixation process which hooks up consciousness to brains happens for, and I quote, "whatever reason." A model would consist of formulating for us a possible reason for this mysterious behavior (not to mention telling us what the fixation actually consists of). Even if it's only a story about how consciousness is a parasite that feeds on brainwaves and thus gravitates towards brains, at least give us something resembling an actual model.

This is a good point in the debate to introduce one of the classic problems of dualism: the interaction problem. Dualists tell us that consciousness is a completely different sort of 'stuff' from the physical world. They tell us that our minds are the mental stuff and our bodies are the physical stuff. If these are such irrevocably different sorts of realms of stuff, it would be logical to conclude that a mind and a body can never interact with each other. Understandably, very few dualists (Locke being the exception) are eager to embrace the idea that our consciousness never influences our bodies and our bodies never influence our consciousness, yet they still have no explanation for how it could work. If there is this physical world separate from our consciousness that the dualist tells us there is, then we cannot possibly be conscious of it.

To borrow the metaphor which my intro to philosophy professor used, interaction between the mind and body in dualism is like trying to throw a softball at the gross national product. Where do we aim? If they're such completely distinct sorts of things which exist in their own separate realms, how can they possibly interact? Put simply, they can't. No physical signal could be emitted which would take a message to consciousness (which direction would the signal fire in?). No conscious thought could ever grab your arm and lift it up, nor send a signal to the brain to do so since the brain is clearly a physical object. This sort of telekinesis is not possible. Divine intervention is apparently required, and is in fact what many dualists end up resorting to. (Note that even *with* divine intervention, God would have to be creating some sort of physical signal for the brain to tell it what consciousness wants... this would violate the law of conservation of matter/energy, and has obviously never been observed.) Indeed, the interaction problem demonstrates that if dualism were true the only rational conclusion to draw would be that *everyone* is a zombie, for there is no way a consciousness could associate itself with a brain. Yet we have the self as a clear counterexample to that, indicating that dualism is not true.

It's also worth noting how the dualist is attempting to give consciousness the same status as the physical world, as being an almost tangible realm of 'stuff.' It is as though they're creating a second physical world and hiding it behind a curtain. For example, in his most recent post AKG tells us that there has to be a conscious entity "affixed" to the brain in order to create qualia. Do we need super glue for this, or can we just duct tape a softball to the gross national product? Even the very talk of affixing smacks of treating the realm of consciousness as being physical. We can easily note that there's time in consciousness, and if there are separate individual instances of consciousness which are lined up and "affixed" to brains then apparently there is space in this supposedly non-physical realm as well. (Once you have multiple points on a graph being defined as unique, which we have here by having separate minds uniquely affixed to different brains, you have space.) This was Gilbert Ryle's point when he described dualism as the dogma of the ghost in the machine ("Descartes' Myth" from [The Concept of Mind](#) I believe)... Ryle points out that the ghost, too, is being constructed as mechanistic. The dualist creates a mechanistic ghost for a machine and pretends that this has somehow explained something.

As has been demonstrated, any attempt to say that consciousness and the physical world objectively exist in separate realms (hence any objective [aka non-Kantian] dualism) ends up collapsing into the same problems. They only create pseudo-physical realms, they provide no explanation for the association of minds with bodies, they make it logically impossible for the consciousness to be aware of the physical body or the body to respond to the consciousness.

Clearly, AKG's main argument is that because consciousness is non-physical it is "unacceptable" to use our senses to detect it. This argument is either a confused one or rests on the presumption of dualism. The sense in which qualia/consciousness is nonphysical is the same sense in which *everything* is nonphysical. Qualia is not a magical land *beyond* the physical -- it is that which a certain type of physical structure represents. The way in which we use our senses to determine if consciousness exists is not in any fundamental way different from the way we use our senses to determine if the chair we're sitting on exists. The chair-in-itself is nonphysical as well in the same limited sense. The difference between the chair-in-itself and the brain-in-itself is that we cannot make an analogy between ourselves and the chair. Going back to Nagel, we cannot imagine what it is like to be a chair. Since chairs do not have any sort of interactive feedback patterns in them indicative of the sort of information processing capabilities that we find in brains, we can rationally conclude that there isn't anything which it is like to be a chair -- the chair-in-itself exists, but there's nothing that it's like to be the chair since what it consists of is something other than thought. Note that we are once again using a physical criteria here to determine that there is nothing that it is like to be a chair, just as we use a physical criteria to determine that there is something which it is like to be a bat (even if we can't imagine what it is like). I must ask AKG if he believes it is possible for chairs, rocks, bricks, baseball bats and various other such objects to be conscious. It would seem that he does, since he has already told us that it is "unacceptable" to use our senses to determine that the chair lacks consciousness.

AKG has, not for the first time, asserted that my position is harder to defend than his and used this as though it were an argument. He seems to be returning to the global skepticism he hinted at in his original post, telling us that we should believe anything is possible. It is in one sense possible that  $1+1=5$ , as everyone could simply be suffering from mass hypnosis making them believe it equals 2. It is not, however, *rational* to believe that  $1+1=5$  (see the discussion of the self-defeating nature of global skepticism in my last post). Neither is it rational to believe that zombies are possible.

We seem to be in agreement that for zombies to be possible requires that some form of dualism be true. Without dualism, zombies become metaphysically incoherent. With this in mind AKG must demonstrate that dualism is a reasonable position to hold. With his mentions of the difficulty of establishing impossibility, he appears to be trying to fall back on saying that it is possible that dualism is in some vague yet to be defined sense *possible*. This is no better than saying it's possible that  $1+1=5$ , until he can provide us with good positive reasons to believe in dualism and he can answer the arguments I have made for the irrationality of dualism. He must establish that dualism is a reasonable hypothesis.

Remember that the peculiarity of global skeptical hypotheses (GSHs) is that assertions made in them are just as impossible to provide evidence for as against. I remind AKG that the topic of this debate is "whether

zombies are possible." If he hides behind global skepticism he can say that zombies are not technically *impossible* -- but the GSH is constructed so that it lies outside of the bounds of possibility and impossibility, so he *cannot* rationally make the positive assertion that zombies are possible. AKG must establish that it is rational to present the assertion that zombies are possible.

By Paul Knierim

Paul's first charge is unfair. This is not the debate for me to propose a detailed model describing the mind body relationship, and I feel no need to give you just any model, it would only be speculation. How, or even if the consciousness affects the mind, how the consciousness affixes to the brain, etc, are irrelevant. Indeed, they are debatable issues but they are **not** matters in this debate. The existence of free will was not questioned in the previous debate, although surely it could have. However, the question was whether free will could be compatible with determinism.

I did provide a model, in fact. There are various models: some suggest the mind is simply an "illusion" which is in fact simply just brain patterns, some suggest that the brain itself is simply a perception apprehended by the mind, I suggest that the mind and brain are separate entities, but in my case the mind is affixed to the brain, resulting in the mind experiencing qualia corresponding to the electrical signals in the brain. Why does this fixation occur? Who cares? If you really think this is worth debating, it is way off topic and please start another thread on this question, but this debate is not the place for it.

Paul also claims that it is illogical to conclude that the physical brain cannot interact with the non-physical mind. Why this is, I cannot fathom. One could conclude that it is illogical to conclude that action at a distance exists. Why do masses just attract, without being in contact with each other? Was Newton off his rocker in suggesting such a phenomenon? Not at all. Masses simply interact in that manner, despite the fact that we have the intuitive feeling that objects need to be in contact to apply a force to one another. It is indeed possible that if the consciousness does affect the mind, that it does despite the fact that they are of different natures, even though this goes against Paul's intuition.

Paul continued with a few arguments however some seemed either irrelevant, unsubstantial (there was no good reason given to think believing  $2=5$  was similar to believing in zombies), or refuted by comments above or in previous posts. He then goes on to say that I have provided some yet to be defined definition of possible. I thought I had been clear on this. I claimed that my position was logically consistent and not in contradiction of scientific facts. This was my definition of possible, I thought that was clear. However, if not, I believe I've shown my position consistent and not in contradiction of science.

I suppose I can do little but repeat myself, but that's what I must do. What is a zombie? Something that **acts like us**, has the same **physiology as us**, and **does not experience qualia**. Can something act like us without experiencing qualia? Sure. The brain is sufficiently complex to suggest that it is possible, if not true, that the brain determines our behaviour. Not that which experiences qualia, but the brain which "experiences" electrical and chemical signals. Our actions are possibly results of very complex chemical and physical reactions in the brain. Can something have our physiology without experiencing qualia? Is that which sees redness rather than simply reacts with electrical signals something necessarily resulting from the brain? I can see no reason to believe so.

Now the above does rest on the idea that qualia are different from brain signals, the consciousness/mind/observer is different from the brain, that redness is different from a specific brain pattern, etc. Is this so inconceivable? Is it even possible to conceive otherwise? Paul suggests that consciousness is what it is like to be the brain, and redness is what it is like to be a certain brain pattern. I think this is entirely meaningless. He says specifically, "[q]ualia is what it's like to be the brain." So if I were a brain, I'd be qualia? I don't think that's it. I think there needs some clarification on what is meant by, "what it is like to be." Does it mean that if X is what it's like to be Y, then X is what Y experiences. Are qualia what the brain experience? Again, that's wrong. The brain doesn't experience redness, it reacts to

brain signals. Whatever "what it's like to be" is supposed to mean, it is in fact meaningless, and whatever it could mean, I don't know how qualia could be what it is like to be the brain.

EDIT: Something I was going to put in but forgot -- it is not even necessary for me to explain how a non-physical entity would interact with the brain and cause the body to do things. You assert some sort of monism, and generally show a lot of faith in scientific facts, etc. You do claim that we have free will, but whatever that is, would you not accept that our behaviour is entirely determined by the physical processes of our brain? What else would it be? If it were something else, you'd be ascribing to some sort of dualism which you don't, so you accept that the brain is responsible for all of our actions, i.e. the physical/chemical/processes determine our actions. This is what I'm proposing, and since I imagine it is acceptable for you, you can see that we can't infer that something other than the brain (i.e. the consciousness) exists from behaviour. And of course, just as redness is something other than the brain signals, and is not physical, consciousness is something other than the brain, and not physical. The difference is I am willing to admit that brain signals cause qualia (in the presence of the conscious entity) but see no reason to believe that the brain necessarily causes the actual conscious entity (the fact that I have I such case is by no means satisfactory).

By AKG

AKG begins his post by declaring that it is unfair to ask him to provide a coherent, reasonable explanation of how something he insists can happen would actually be possible. Good heavens, what was I thinking? We must just take it on faith that any absurdity a person can utter can happen. He goes on to state that he considers anything and everything to be a model, and that keeping it logically possible is by no means a requirement or even a good recommendation to take under advisement when creating your model. No purpose to debating semantics, so I will simply note that the way I use the word "model" requires that it be an illustration of a possible explanation for something.

Let's spell things out once more:

(Although I've already supported these premises elsewhere, we'll treat them as premises here for illustration purposes.)

Premise 1: If dualism were true, then everyone would be a zombie due to the interaction problem.

Premise 2: If the self exists as both mind and matter, then not everyone is a zombie.

Premise 3: The self exists as both mind and matter.

Premise 4. If dualism is false, zombies are not possible.

5: Via modus ponens on 2 and 3 we have that not everyone is a zombie.

6: Via modus tollens on 1 and 5 we have that dualism is not true.

7. Via 4 and 6, we conclude that zombies are not possible.

Rewritten in a purely symbolic form (P=dualism, Q = everyone zombies, R = self as mind and matter, S = zombies are not possible):

1.  $P \rightarrow Q$

2.  $R \rightarrow \sim Q$

3. R

4.  $\sim P \rightarrow S$

5.  $\sim Q$  (MP 2,3)

6.  $\sim P$  (MT 1,5)

7. S (MP 4,6)

QED. Since AKG seems to think it is unfair to ask him to debate dualism here, we should apparently just accept that zombies are not possible. If AKG does not admit that he is conceding, he must give us a

*rational* model of interaction to explain why it would not be clearly the most logical to conclude if there are two irrevocably separate realms -- so separate that we are told that it would be completely illogical to place a softball in the mental realm because it's supposed to not be that 'sort of stuff' -- that it would follow that everyone human body be a zombie (and every mind a disembodied spirit).

The zombie scenario was in fact invented to 'prove' dualism. While I will try to keep outside references to a minimum, I must refer AKG to the Stanford Encyclopedia of Philosophy's entry on zombies:

<http://plato.stanford.edu/entries/zombies/http://plato.stanford.edu/entries/zombies/>

The encyclopedia notes, "It is argued that if zombies are so much as a bare possibility, then physicalism is false and some kind of dualism must be accepted." Pretending that the issues are unrelated is not an option.

Note that some physicalists like to say that we are all zombies. I certainly disagree, but I won't detail my argument unless AKG wishes to argue that everyone is a zombie. What matters here is that there is one and only one theory of mind which makes it possible for *some* people to be zombies and others not: dualism. Without dualism, it is absolutely impossible to have a mix of zombies and non-zombies walking around the world. If dualism is false, it is very clear that this means there are no people with more or less consciousness than others.

The "I can't fathom it" response to the interaction problem seems to indicate that AKG is treating dualism like we would treat differing sorts of things in our unified physical world. The obvious and immensely important difference he ignores is the fact that everything in the physical world is related in the sense of space and time.

How far is your consciousness from the floor, AKG? What's the diameter of it? When you tell us that your consciousness is affixed to your brain, how much of the surface area of your brain does it cover? How many millimeters of air space are there between the consciousness and brain? Sure, right, there ain't no problem at all to signal between them... it's somehow just like signaling across space he says, and he can't fathom anyone not agreeing. Fine... all I ask is that AKG take a moment to actually *tell us*, then, how this works. This is why I asked him for a model, meaning an explanatory model that actually gives a possible example of *how* something could happen. Yet he has not done so and has indicated that he will not be doing so. I hope that in his concluding post he will finally try to provide a solid positive argument for dualism which gives us a clear picture of how it can actually logically work, and I look forward to it.

Next, AKG takes an embarrassing foray into science, which apparently isn't his field. He declares that there is no reason to think it's illogical to throw a softball at the gross national product, because Newton said we can do magical action at a distance. Of course it's a nonsequitor to begin with since action at a distance and action between metaphysically separate realms is a different degree of nonsense, but we'll set that aside. More interestingly, AKG's Newtonianism is a nice illustration of a system (Newtonian gravity) that sounded nice a surface glance but died due to unresolvable philosophical problems. Newton's action at a distance was indeed utterly wrong -- Einstein rejected and replaced it by the early 20th century, and Eddington experimentally proved Newton's theory of gravity wrong with his measurements of the solar eclipse in 1919. Newton was of course trying to do science more than philosophy, but action at a distance has gone down in history as Newton's philosophical blunder that broke his theory of gravity, and the death knell of the materialist philosophy which had flourished based on Newton for a time. Clearly Leibniz was shown the winner of their debates in the end.

"Masses simply interact in that manner," AKG tells us, throwing away a century of science as he strolls onward and apparently showing us the sort of blind faith in irrationality and illogic that we would need in order to accept dualism. Sorry, AKG, but gravity simply does not work in the Newtonian action at a distance way, and the rest of the world has known this for nearly a century. I await AKG's refutation of general relativity and his clear reasons he's sure to provide us for throwing out Einstein's law of gravitation -- a law that was incidentally inspired by the clear philosophical problems of Newton, and in particular Mach's objections. Einstein credited Mach's philosophy for helping him see Newton's mistakes. While I don't have space here to explain the law of gravitation, the e-text of Einstein's "Relativity: The Special and General Theory" is linked in the links section of this website.

Finally, AKG makes an odd argument that free will implies dualism. He seems to present us with a choice of either hard-core eliminative materialism and hard determinism on one hand, or dualism as the only other choice. This is so patently absurd that his 'inference' drawn from it would be irrelevant regardless, but every eliminative materialist hard determinist in the world says that our observations and internal reflection will indicate to us that we feel we have free will. Hence we cannot derive anything from observation of our feeling of freedom when both systems predict the exact same observations to result.

Note the lack of any response from AKG to the point that dualism is self-defeating because it simply creates a mechanistic ghost. Due to word count limitations I didn't fully expand on that in my previous post, so I'll continue here. The physical nature the dualist ascribes to consciousness arises because the dualist is implicitly treating the consciousness as an object, which could in theory be observed with magical mental eyes (in fact, some say these eyes do exist in the form of psychic powers). The dualist has to treat it like this, in order to call consciousness an objectively separate metaphysical realm. Dualism can only arise from implicitly treating consciousness as a third person perspective. The fatal flaw of the dualist is a blindness to perspective: the dualist takes something which is the very definition of the first person perspective (consciousness) and shoves it into third person ontology of objects, assigning it pseudo-locations and movements and methods of affixation.

Even as we enter the final round of the debate, AKG still has yet to provide a coherent explanation of how dualism can be possible. It necessarily follows from this that he has not shown how zombies can be possible.

By Paul Knierim

The first half of Paul's post relies on some sort of "interaction problem." How can two things of differing metaphysical nature (the mind and the brain) interact? However, Paul has taken no effort to show that it is impossible. I can show that it is possible, but that is not very difficult and indeed not at all relevant. I say this because I don't believe we have any known and tested laws of metaphysics as we do for physics. I don't have to worry about breaking these laws so it is not too hard to dream up any logically consistent system. Moreover, that which observes, the consciousness, is itself unobservable, so we cannot know the details of the nature of the mind/self/observer. Because of this, I have even fewer restrictions.

I maintain that Paul's demand is unfair because it is irrelevant and avoids the main issue, asking me to explain things that would only be speculation. Brain patterns produce qualia if there is a mind "affixed" to the brain. How does this happen? The mind is some sort of causal agent which allows it to happen. How does fixation occur? Of course there's no super glue or duct tape involved, how would you like me to answer? There is some metaphysical bond. If it's not glue and not duct tape, what is it? I know it's not glue or duct tape, but how can you expect me to tell you what it is. Even if I said something, you wouldn't know what that is. It's really an unfair question. So what is the nature of that metaphysical bond, and what is the mechanism which allows for the fixation? You may as well ask me how many 5-dimensional being it takes to screw in a 4-dimensional light bulb. I could give any number without violating a known law of 5-dimensional beings. I could even provide some sort of speculative reasoning because we can't really comprehend or observe a 5-dimensional being or a 4-dimensional light bulb, so just about anything goes. The mind cannot observe the mind in itself, and so we cannot know much about the nature of the mind in itself, so asking me to give a detailed account of how one interacts with the other is absurd.

We know that glue allows us to attach to physical things. Why must it be that we cannot attach a physical thing to a non-physical thing? What rule says that? All we have to do is understand that we have to go outside our regular intuition and consider things in a different context, and it's not so hard to imagine. In truth, the consciousness doesn't have to be something outside of space and time, it is simply some entity distinct from the brain. Some people have even proposed that it's just a different type of matter or energy, one we haven't dealt with, just like 5th and 6th spatial dimensions which sound strange and sci-fi-y but can certainly exist simply without our ability to understand them.

Moreover, to claim that it just can't happen is absurd. Why is it that it just can't happen? How is this even an objection? What reason is there to believe there even is an interaction problem (other than Paul's intuition)? Now Paul correctly points out that the argument about gravity and action-at-a-distance is outdated. However, I can offer a more modern and in fact more convincing argument. Consider the [EPR Paradox](#)

*Today, most physicists agree that local hidden variable theories are untenable and that **the principle of locality does not hold**. Therefore, the EPR paradox would only be a paradox because **our physical intuition does not correspond to physical reality**.* (excerpted from link above; emphasis mine)

The EPR paradox, for those who don't care to read the article, is about the instantaneous transfer of information (faster than light speed) between two particles. The very principle of locality which Einstein held to and used as motivation to show Newton's action-at-a-distance theory of gravity wrong turns out to be **false** as demonstrated by the EPR paradox. It is simply Paul's intuition that prevents him from accepting that reality interacts in certain ways. With that out of the way, Paul still demands that I explain to him how this interaction is then going on. If I pick up a box and shake it and here something inside, I know something is inside. Paul is now asking me to tell him what's inside without looking or feeling around inside. How the mind works is beyond me, I don't see why I must be expected to answer this. [Paul is just moving the goalposts.](#)

The remainder of the post by Paul is vacuous (stating things are absurd without justification), repetitive (asking me to explain what's in the box I can't see inside), or pointless (saying that the dualist treats consciousness as something different from the physical, without showing the problem in this). This was explained long ago. Redness is distinct from anything in the brain. It is not the brain, or the brain signals, or anything like that. Although science suggests it is caused by the brain, (refer to the example of prodding the brain resulting in a person claiming to see something different), they are indeed different things. The consciousness, I propose, is a different thing from the brain, or "what it is like to be the brain," depending on what that even means. This is not so hard to understand and see as true, if not at least possible. How does the mind affix itself to the brain? I don't know Paul, what's in my drawer right now? All you can do is speculate, and that's all I can do now as well. However, you know more about the nature of drawers than I do about the nature of the mind, so it is even more absurd for you to expect any sort of answer here. It is not that I am unwilling or afraid to give an answer, but one would have to be so vague and so general, there would be no point. I can't give details because I can't observe the metaphysical nature of the mind. So what good is an explanation without details?

So can something behave like me and have no consciousness? Well if the consciousness is metaphysically different from the brain, then sure. The brain, we agree, determines behaviour. Paul must necessarily take this position because for him there is nothing other than the brain. If there is a brain, must there be a consciousness (assuming metaphysical differences between the two)? Not necessarily, Paul would have to prove that there is a necessary entailment of there being a consciousness if there is a brain. So the zombie problem is solved, with the only thing in question is dualism, i.e. are the brain and consciousness indeed metaphysically different? Redness and brain signals are different. There is a notable difference between "I perceive red" and "the brain interprets brain signals." It's more than just a semantic difference, there are distinct differences between the perceiver and the brain, as well as between redness and brain signals -- these are metaphysical differences. You can't mail me redness, or the image you perceive right now of your desk and computer, etc., but you can mail me a piece of your brain. If the consciousness is different metaphysically from the brain, can they interact, i.e. can the consciousness affix itself to the brain? Of course, why not? Can I explain how this happens? Not in any relevant manner, if you insist I make one up, well, I decline because it's a pointless waste of time. I also can't explain how it is that the universe exists, so do I have good reason to believe that it is not possible that the universe exists? It is indeed possible that the universe exists (and depending on how you look at the question, it is necessary). I can't explain why the law of conservation of energy exists. I can't explain why light speed is constant. I can't explain why certain physical constants are what they are, but do I have reason, then, to believe that they are not? Do I have good reason to believe that the mind cannot interact with the brain if they're metaphysically different? Not a chance. Maybe Paul's intuition would get in the way, but the EPR paradox suggests that he needs to second

guess his intuition.

Really, this is not a difficult position to understand. It is not hard to see why zombies are possible, and why dualism is possible (if not necessary). It is plain that the image of the red apple on the desk we see is different in nature from the physical desk and apple in themselves, the light which reflects off of them, or the electrical signals in our brain. It is not difficult to recognize that consciousness is distinct, and that it can affix itself to a brain, and a number of other facts. All Paul has to do is take a minute to realize he's letting his intuition stand in his way. Hopefully all the readers can analyze my model logically, notice its consistency and its non-contradiction to science, and realize the validity and soundness of the position without the analysis being coloured by their intuitions.

By AKG

Note that I'll ignore AKG's talk of mailing redness and qualia and the like due to word count limitations. It's evident that his claim that qualia implies dualism is false, since other theories explain qualia much better (not hard, since as noted dualism doesn't actually explain anything, it merely waves a wand to declare that there must be an explanation somewhere).

AKG makes an excellent point in the spots where he says we must not be slaves of intuition. Intuition can be an effective check to make you reevaluate it if the position you're holding contradicts with everything that seems obvious (hence the purpose of philosophical 'intuition pumps'), but it is not a law. Unfortunately he fails to heed his own advice, since he has again and again based his arguments on his intuition that there must be a magical solution out there somewhere for dualism since he happens to intuit that feels right to him. How his intuition makes it seem to him that the difference between the appearance of something and the thing in itself implies dualism, as he says at the end of his post, is certainly beyond me – particularly as that very distinction is what Kant used to demonstrate the impossibility of objective dualism in the CPR. Most likely he is confusing empirical dualism with objective dualism, which I suppose would be an intuitive thing to do, yet seriously incorrect. I suggest AKG cease relying on his intuition.

AKG has now outright proposed that there are magical metaphysical laws which will make whatever he wants possible. I will agree with him that we don't have any known and tested laws of metaphysics -- that, obviously, is because it's utter nonsense to suggest that there could be such things. The kind of substance-metaphysics he imagines is impossible to logically defend for a simple reason: it denies all evidence of itself. AKG proposes that because we don't have a scientifically validated cookbook for witchcraft, it somehow follows that he need not provide any evidence for saying that his spells are possible.

All of AKG's arguments could just as easily be applied to argue that it is possible that  $1+1=5$ . I hereby propose a model in which there is a magical metaphysical law (which I need not explain the nature or functioning of, because it's metaphysics and metaphysics has become our code word for any kind of ad hoc magic which we don't wish to explain) which makes  $1+1$  equal 5. To the "sum of two instances of a number can't equal 5 times the number" counterargument, I say that in the 5th dimension we don't know how addition might work. Now, to all objections anyone may try to bring against my beautiful model I will declare that you cannot prove that my model is not possible because my model is a metaphysical law, and there are no set proofs of which metaphysical laws are valid or invalid-- meaning I can say my cherished law is a possibly valid metaphysical law. Hence, according to AKG's version of logic  $1+1$  may equal 5. Kids, I suggest you don't try this on your math test. (Now for the next trick he can just as easily demonstrate that colorless green ideas may possibly sleep furiously.)

A point-by-point breakdown:

- It is a necessary consequence of the AKG's position on the invalidity of physical evidence that we would never have any reason at all to think anyone conscious [short of psychic powers, which he did not argue for]. We would be compelled to become either solipsists (dismissing the whole universe as imaginary) or eliminative materialists (dismissing consciousness and qualia as

- imaginary). Neither solipsists nor eliminative materialists, incidentally, can claim that some people have consciousness and others not – so they cannot believe in the possibility of zombies.
- AKG rejects the common sense ways we do every day evaluate people for consciousness, yet at the very same time he argues that we are *usually* correct in concluding that people are conscious or rocks are not conscious. He's being internally inconsistent. If he followed his arguments through rationally, he could not believe that anyone is conscious unless he claims to be using psychic powers to directly see people's consciousness without going through physical means.
  - He has argued for the unacceptability of using our senses to determine if consciousness exists, yet has failed to counter to point that this forces him to say we can't use our senses to determine if anything at all exists. According to AKG's arguments we cannot conclude from physical examination of something that the thing is actually there -- we can only conclude that we have sensations, apparently, since the sensations are all that we are in direct possession of. Here he is compelled by another of his own arguments to be a solipsist, and solipsism makes zombies impossible.
  - He bases his arguments on shaky attempts at interpreting science. Of course, the fact that scientific theories have rested on declaring that basically magical things happen is exactly what has in every case lead to the eventual downfall of the scientific theory proposing it. Science is always a work in progress... we settle for irrational placeholders for a period of time, but eventually we find the real explanation to replace it. He asserts later that the EPR paradox somehow supports Newton, which is rather an odd claim. EPR leads us directly to Bell's theorem. Bell gives us the options of non-locality, idealism or throwing out logic. All scientists choose one of the first two. Idealism directly contradicts to dualism and zombies. Non-locality indicates the interconnectedness of the physical world, not mysterious action at a distance... the speed of light is the limit for transmission of information, non-locality cannot transmit information, cannot be involved in actions, so clearly does not help AKG.
  - It is absolutely impossible to provide a counterexample to the rule of brain patterns meaning consciousness, and AKG has not debated this. Surely we can see that there is no meaning to claiming that something is not true if there is no *possibility* of a theoretical counterexample. It's the same as saying that there's an invisible undetectable pink elephant standing on someone's head. Suggesting that zombies are possible is just as irrational as suggesting that the undetectable pink elephant is possible.
  - AKG has failed to demonstrate how dualism, which is a necessary precondition for zombies, can be possible. He has not given us a rational model of how it can work. His model is only a model in the sense that "1+1=5, because 1 could act like 3 and 2 depending on what the weather is like in the 10th dimension" is a model. He has not solved the interaction problem, or even made a reasonable attempt to address it.
  - My opponent admits that he can't tell you how anyone might go about creating a plan to throw a softball at the gross national product, yet he continues to baselessly assert that it is possible.
  - AKG provides no argument against the fact that if dualism is not accepted zombies are impossible, and yet also gives no solid reason for accepting dualism. Dualism is a precondition for the possibility of zombies... without dualism being established, zombies are clearly not possible.

From all of this I must conclude that to propose the possibility of zombies is irrational.

In practice, what the dualist is trying to do is use zombies to prove dualism even though the possibility of zombies clearly requires that dualism be true in the first place... in other words, all the dualist is doing with the zombie matter is begging the question.

Dualism tries to dredge up the problem of other minds in the disguise of zombies and pretend that it is unsolvable. It admits to having no evidence, it admits to having no solid theoretical grounding... and of all things, it dares to appeal to intuition as its sole backing even though it stands in contradiction to our intuitive sense we demonstrate every day, our common sense ability to realize that a person's brain has quite a lot to do with whether they're conscious.

We ask a person if he believes there may be zombies, and he says yes. We ask him if he's seen any evidence, and he tells us it's absolutely impossible to ever see evidence for it because no physical investigation can determine such a thing... physical investigation has been declared an invalid method of inquiry. We decide then that he must have a really good theory behind it if he believes it despite the impossibility of evidence, so we ask him how it works. He states that he has no idea how it works, it's something far beyond his grasp. We can surely see that the person's belief is utterly irrational. Perhaps you can believe something you can't find any evidence for as long as you have a really good theoretical concept behind it, but if you don't have a theory either then your belief is clearly not justified.

I thank AKG for the enjoyable debate.

By Paul Knierim