MARIANNE DRABEK

What's the problem with free will?

REVIEW

The debate about the existence of free will is often referred to as "the free will problem". However, this essay intends to illustrate that "the free will problem" may not exist after all but that discussions about free will suffer from a multitude of problems. In particular, definitions and theories lack consensus about the nature of free will and are often too general. Support for arguments is often provided by citing neuroscientific studies despite the fact that these suffer from methodological limitations and allow multiple interpretations depending on the very definition of free will. Furthermore there is the risk of undermining the belief in moral responsibility in society by premature public discussions.

"The" problem of free will is therefore, according to this essay, that inter- and intra-disciplinary teamwork in this debate is not optimal yet. It is concluded that a consensus on a definition about free will is a prerequisite for advances in this field of science.

Keywords: free will, free won't, moral responsibility, neuroscience, philosophy

Marianne Drabek; research master student Cognitive Neuroscience
Maastricht University, Maastricht, the Netherlands

m.drabek@student.maastrichtuniversity.nl

INTRODUCTION

The existence of free will has been debated amongst philosophers for decades. This debate has drastic implications for every individual as it is most natural for humans to believe and rely on the assumption that everyone is the master of one's own actions. Many (if not most) individuals would feel deeply uncomfortable if this is not the case since it would mean that they perform actions and express wishes that they do not generate consciously, not unlike a puppet.

At the same time the nonexistence of free will is often claimed to be a threat to society and to require fundamental changes in its penal system since wrongdoers cannot be held accountable for their actions if they are not masters of their will. 'Unsoundness of mind' is an important concept in the legal system but denying the existence of free will altogether implies that nobody could be held responsible or punishable for anything. The mere thought that crimes like murder or rape would not even be attempted to be prosecuted is dreadful.

The tremendous potential for conflict that this topic holds for society is mirrored by the eagerness to discuss this matter: the search term "free will" yields 19,000,000 hits in google, 3,590,000 hits in google books, 263,000 hits in google scholar, and 4,741 hits in ScienceDirect (19.04.2013). While the chosen databases are only examples and these numbers are only approximate, the order of hits is informative enough to imagine the dimensions of this debate.

In the last 30 years this debate tightened as neuroscientific experiments are often cited from critics as proof for the claim of free will being an illusion. Yet, what exactly is being cited and what did these neuroscientific experiments show?

This paper aims to elucidate some problems associated with the *debate* about free will and tries to assess their potential. Is the non-existence of free will supported empirically? What does the debate mean to society? And more importantly: What is being debated, in other words: what exactly is free will? Some answers will be given in this essay, in particular by showing that the answer to the last question is far from easy.

The first section reviews some definitions of free will and sketches some considerations about defining free will. This is followed by reviewing neuroscientific experiments with regard to free will and their limitations. The third section discusses possible issues for society arising from the current debate. Lastly, a summary with conclusions completes the essay.

WHAT EXACTLY IS FREE WILL? PROBLEMS WITH DEFINITIONS AND THEORIES

While the meaning of free will seems intuitive for most humans, it is not easy to give an exact definition, as the following section will show. Philosophers often discuss free will with regard to moral responsibility and its existence, mostly by referring to common views including (in)compatibilism, (in)determinism, libertarianism and subtypes thereof (for an overview see Pockett, 2007). Yet, few philosophers describe the *underlying nature* of free will. However, this is important, because definitions determine how arguments are used and what exactly it is that is being discussed; many times this seems not to be taken into account. Neither intending to summarize all possible opinions nor to argue for or against any perspective, a few accounts on free will are described in this section to demonstrate how different assumptions about the nature of free will influence the overall debate.

Problems due to different assumptions about free will

Some philosophers, the most famous being Descartes (see The sixth Meditation, translated by Moriarty, 2008), separate free will from matter and describe it as a metaphysical force from which thoughts and actions originate. This dualistic view is clearly fairly mysterious, neither possible to prove, nor easy to disprove completely. Others, like Wegner, describe free will as a feeling, sensation or emotion, "not unlike happiness or sadness..." (Wegner, 2004, p.658). This way, Wegner argues, acts are attributed to the self in a post-hoc manner which makes free will illusory. The logic is that if the *thought* is *consistent* with and occurs before the movement and if there is no other "obvious cause for the movement [t]hese features imply causality, that the thought led to the movement" (Hallett, 2007, p.1182).

While he does not distinguish properly between feeling, sensation, emotion, or percept, Gray Hardcastle illustrates the importance of exact wording. She commented that "[o]ur sensation ....may get it wrong once in a while; it may get it wrong lots of times. Nevertheless, the sensation is reflecting something real, as real as our bodies' need for nutrients. The important question is what exactly is that sensation reflecting." (Gray Hardcastle, 2004, p.663). When looking at free will this way, this definition goes beyond mere perception; mistakes may happen but there is a 'real' basis to the sensation. Already such a relatively small difference (compared to viewing free will only as a perception) reshapes arguments about free will.

Yet others, mostly compatibilists like Frankfurt, link free will to decisions or choices and thus ability (for overview of compatibilism see Campbell, 2011). It is important to distinguish this account from the other in that decisions necessarily precede actions, a feature that has important implications for arguments about free will's existence. Already these few examples hint at the degree of variation that a definition on free will can include: In comparison to the meta-physical force account, viewing free will as a percept or emotion offers more loopholes, e.g., that a feeling may be mistaken in certain situations and also explain some of the function that free will may have.

Problems due to missing limitations in definitions

In fact, the inclusion of loopholes and limits make definitions resistant against arguments about exceptional cases and thus make them generally valid (i.e. the definition cannot be proven to be wrong by such arguments because exceptions or limits are specified). Unfortunately, the range of cases to which a definition can be applied is rarely discussed although this is a crucial component: Is it necessarily a (an entirely) conscious process and what process is defined to represent free will: "The earliest initiation of action process" or "the choice or selection of a *specific* action" (Haggard & Libet, 2001, abstract. Bold print in original instead of italics)?

When does free will develop (or does everyone possess it from birth)? Does everyone possess free will (e.g., children, easily manipulable individuals, cognitively impaired or mentally disordered)? Under what circumstances is one capable to exert free will concerning specific decisions (e.g., being in love, exceptional emotional situations, physical conditions like hunger or sleep lack, substance abuse, peer pressure, brain tumors)?

Especially with regard to definitions involving decisions, it may be required to impose certain restrictions otherwise discussion may be strangled by arguments about physical needs (e.g., mild hunger when shopping interfering with choices) and impulsivity. Restricting free will to higher-order processes would help to clarify discussion grounds and avoid debates about ambiguous situations. At the same time, even such a restriction may be too vague because emotions and context influences can never be completely left out of any decision. Thus, it is needed to specify the type of decisions in question and the degree of being informed about the different choices and corresponding consequences.

When free will is described, it is often briefly defined to involve no coercion or constraints. However, decisions may not necessarily reflect the will of an individual because of the limited choice possibilities of a situation. Examples demonstrating that a person's will may not match any of the possibilities in a particular situation are plentiful (even if budget is unlimited, certain desired things may not be available). This makes it questionable that true "no-constraints situations" exist in a nonutopian world and thus precision is needed in such statements. Would it be better to speak of a person's 'wishes' or 'desires' rather than 'decisions'? This could be an entire discussion on its own; the example is only mentioned here to illustrate again the influence of phrasing. Important is, though, that 'free' should not be confounded with freedom.

The previous point was raised about external coercion (or constraining) factors. What about internal factors: Could one not argue that past experience or physical needs already dictate? If so, one could object that a person will never be free because of influences of the past and habits or attitudes that resulted from it (e.g., eating sweet things since childhood); but then exactly how free is 'free'?
One could argue that attitudes are part of the person's personality rather than internal coercion factors since the person may have actively decided to adopt a particular stance (e.g., 'I do not buy this product because I'm vegetarian out of principle....'); yet again these conscious decisions and attitudes may have developed due to past experience. On the other hand, some consistency in a person's decisions is expected (since they are generated by the same person) but at what point is the line crossed to favor determinism?

While certainly leading, a person may still decide differently from what attitude or habits dictate (e.g., '....but this time I'll make an exception'). Is the ability to veto enough to prove that there is a 'real choice'? Some philosophers and scientists think so since they defined free will by quite the opposite, a 'free won't' (e.g., Obhi & Haggard, 2004) instead of an active decision (free will).

This shows that many different cases are needed for thought experiments to develop a waterproof definition instead of a general statement. Certainly many of these questions are bound to be discussed but these jigsaw elements are spread in the literature, not agreed on and often omitted all together.

As a conclusion, free will requires a complex definition because the very nature of free will and possible limitations need to be addressed. This may be tedious but possible if constructive discussions are held. Sufficient time should be taken since

fast and premature definitions lead to misunderstandings, circular arguments and will only hamper progress.

Moreover, the very definition of free will is a basis for experimental rationales and interpretations, as the next section will show, which makes it clear that a consensus on the definition is a prerequisite to advance this field of science with empirical experiments.

PROBLEMS OF NEUROSCIENCE WITH EMPIRICAL INQUIRIES ABOUT FREE WILL

Empirical results are always valid only in the context of a specific definition and with respect to their limitations. While this seems so obvious that it is hardly worth mentioning, limitations of interpretations are often left aside in debates about free will by citing only what are believed to be the main findings. This section explains some of the methodological weaknesses of neuroscientific experiments in the context of free will. In particular, the 'classic Libet experiment' will be dicussed in this section followed by some theoretical and methodological considerations against the common interpretation of this and similar results.

The 'classic Libet experiment'

In almost any introductory part in articles about free will, the so called "Libet experiment" (Libet et al., 1983) is summarized as the first and cornerstone of neuroscientific evidence against free will. Using electronencephalographic (EEG) recordings from subjects, who were told to make a voluntary key press at a time of their choice, Libet and colleagues showed readiness potentials 1 second prior to the execution of the movement. These potentials occurred prior to the subject's own awareness of his or her decision. To show this temporal component of their awareness, Libet's subjects were asked to observe a rotating clock and to note the clock's position when they were aware of their urge to press either of the two keys with either left or right hand. They found that the readiness potential was indicative of the choice and occurred prior to the time of becoming aware of the decision, as indicated by the reported clock's position. This has often been taken as evidence that the subject's choice was determined by something else than one's own free will.

Problems with the interpretation of Libet's findings

Similar experiments have been carried out using single-neuron recordings and fMRI but this experiment is by far the most cited paper in these discussions. Yet, comparatively few articles mention the limitations or doubts against the common interpretation of his data. Interestingly, Libet himself did not claim to have proven the non- existence of free will. On the contrary, he explains that the movements of the subjects could be initiated subconsciously and that while such a decision may be generated without the subject being aware, the mind still possesses a veto power to not execute the act (Libet, 1999). In this respect Libet seems to subscribe to a

"free won't", emphasizing that a definition on free will should not be made light-hearted. However, even if we hold on to the definition of free will rather than the 'will to do otherwise' as it is also referred to, there are more objections against hasty conclusions.

There is the possibility that the activity measured (either by EEG or fMRI or electrophysiologically) *reflects preparatory activities for the decision making or even the decision process itself* (Trevena &Miller, 2002). This process may again consist of both subconscious and conscious components, all of which makes temporal data (e.g., by EEG) uninterpretable in terms of what activity represents free will or what is predictive of free will.

In fact, finding biological components or mechanisms of a decision via EEG or fMRI or single-cell recordings does not even necessarily mean that dualistic believes need to be abandoned because it is known (e.g., from neurofeedback) that the *mind is capable of influencing brain activity* (Stier, 2011, p. 990. translation). Does the mind influence brain activity or is it the other way round? Due to the possible entanglement of the various (sub)conscious components in a complex decision process, this seems to resemble a hen-egg question. Moreover while results may be found that explain certain features, Roskies argues that neuroscience is not in the position to *prove that free will is reduced to [one] mechanism only* (Roskies, 2006, p.421).

This shows that there are already sufficient extensive theoretical considerations which should at least caution against absolute claims (that free will is an illusion) without the need to refer to methodological weaknesses; a selection of these will nevertheless be mentioned in the following.

One potential methodological weakness is the choice of tasks in these experiments. As Haggard notes, participants are usually asked to make decisions about personally irrelevant or meaningless choices, such as right or left key presses. Thus, there is "generally no reason or value that motivates the participant to choose one action over another" (Haggard, 2008, p. 934), which may mean that there was no need for the participant to have any free will in these moments. It follows that tasks may need to involve higher-cognitive decisions; a point that is also relevant when reviewing animal studies on free will. Nevertheless, even when the tasks get more complex or closer to everyday life, the argument may remain that the situation is a laboratory-one and therefore too well controlled by the subject (Stier, 2011), who, in addition, may make (unconscious) efforts to please experimenters.

Further, it is questionable whether the temporal components measured are actually reflecting meaningful temporal order. Trevena and Miller who conducted an experiment similar to Libet's, acknowledge that averaging over trials may cause *smearing effects in readiness potentials* which may completely misrepresent temporal order (Trevena, & Miller, 2002). Further inaccuracy may be introduced by *perception difficulties,* i.e., delay which is introduced when the person indicates the time point of the decision (Trevena & Miller, 2002). Similarly, there has been research, showing that the "subjective present is actually slightly in the real past" (Hallett, 2007, p.4).

On top of that, it has to be noted that the temporal resolution of the methods

to study brain activity may not be close enough to the resolution needed. This is sure for fMRI which has a temporal resolution in the order of several seconds. However, despite the fact that the temporal resolution of EEG in miliseconds is fairly good one cannot be certain that it suffices for absolute claims (Stier, 2011).

As it stands, caution is recommended when using neuroscientific data in arguments about free will. For fairness it should be mentioned that there is also empirical work aiming to demonstrate the existence of free will; it suffers from the same limitations, though: definitions, theoretical considerations and methodological weaknesses (review by Stier, 2011). The best methods currently available may not provide satisfactory quality for the problem studied. Perhaps it is even impossible to create an experimental design that is complex enough and immune to methodological inaccuracies. Is neuroscience able to advance the free will discussions at all? Even though neuroscience did not achieve reliable evidence for or against free will yet, this research has already moved a mountain by initiating interdisciplinary talk. Hopefully this will lead to refined definitions and theories at the least. The next section will briefly show what the debate means to society until the debate has come to a consensus.

PROBLEMS EVOLVING FOR SOCIETY FROM FREE-WILL DISCUSSIONS

Without a doubt, the discussion about free will may have drastic implications not only for the common justice system but also for every individual. While this problem has the potential to be tremendous, as outlined in the introduction, several researchers and philosophers claim that it may not be a problem after all- at least not for anyone who is not a dualist (Hallett, 2007). Hallet for example argues that "[a] person's brain is clearly fully responsible, and always responsible, for the person's behavior" (Hallett, 2007, p.1189). Adherents of a 'free won't' could also argue a person could have used their veto power to prevent the action from execution (Libet, 1999), which would still make them responsible. Again, this discussion depends to some degree on a definition, but regardless of the exact arguments there seems to be some agreement that the legal system should not be changed. If so, is the conflict solved then?

The answer is no, since there may still exist a problem for the society arising from the discussions about free will and determinism. When Vohns and Schooler primed participants with texts on determinism, their participants cheated considerably more often and thus were led to behave immorally (Vohns & Schooler, 2008). Baumeister and colleagues found that subjects behave more aggressive and less social after being primed that free will may not exist (Baumeister et al., 2009).

These and similar data should be a warning to the scientific community that there is the potential for great harm for society when careless phrases and incomplete discussions (until a consensus in science is reached) reach a non-critical audience. Some individuals in such an audience may not assess such discussions correctly and may make use of any excuse to abandon responsibility for misdeeds.

'THE' PROBLEM OF FREE WILL? SUMMARY AND CONCLUSION

This last section summarizes and concludes what "the" problem of free will is (according to the author of this essay). The first section focused on the problems of various definitions and underlying assumptions and illustrated what implications these discrepancies have. Sufficient time should be taken to work on these issues since fast and premature definitions lead to misunderstandings and will only hamper progress. Moreover, the very definition of free will is a basis for experimental rationales, which makes it clear that a consensus on the definition is of outmost importance in order to advance empirically.

The next section described some of the concerns that have been voiced against the common interpretation of Libet's experiments and similar work. There have been considerable theoretical and methodological concerns such that current empirical data are neither suitable to fully support contra nor pro free will arguments at this stage. Such arguments include for example the entanglement of conscious and subconscious components in the decision process or the temporal interpretation of data, or the type of tasks in these experiments. The problem is thus that neuroscience has yet to come up with adequate strategies and methods as to how free will can be studied.

The problem illustrated in the last section concerns problems for society from debates about free will. The illusion of free will (should it be proven) would perhaps not cause problems for legal responsibility. Instead there is an immediate and observable threat in incomplete discussions and premature conclusions, because they may tempt a less critical audience to abandon the concept of moral responsibility, as has been shown to be the case in several studies.

Given that the problem of free will is tightly linked to its definitions, these problems could be expected to be handled after sufficient discussions. However, a prerequisite for this is patience, well-working interdisciplinary teamwork and mutual respect of philosophers and neuroscientists as equal partners in this discussion. In addition, good scientific practice includes that results are always reported and cited with regard to their respective limitations. If this succeeds, eventually, a consensus on a definition will be reached and empirical advances can be built up on this. Until then, necessary caution should accompany discussions to protect society.

Acknowledgments

I want to thank my dad for being a passionate listener and the reviewers for constructive comments.

REFERENCES

Baumeister, R.F., Masicampo, E.J., DeWall, C.N. (2009). Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness. *Personality and Social Psychology Bulletin,* 35(2), 260-268. doi: 10.1177/0146167208327217.

Campbell, J.K (2011). Free Will. Cambridge, *Polity Press.* pp. 86.

Descartes, R. (2008). Meditations on First Philosophy with Selections from the Objections and Replies. A new translation by Michael Moriarty. Oxford World's Classics. New York, *Oxford University Press.*

Gray Hardcastle, V. (2004). The elusive illusion of sensation. [Comment on Wegner]. *Behavioral and Brain Sciences.*

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience, 9*(12), 934-946.

Haggard, P., & Libet, B. (2001). Conscious intention and brain activity. *Journal of Consciousness Studies, 8*(11), 47-63.

Hallett, M. (2007). Volitional control of movement: The physiology of free will. *Clinical Neurophysiology, 118*(6), 1179-1192.

Libet, B. (1999). Do we have free will? *Journal of Consciousness Studies*, 6(8-9), pp. 47–57.

Libet,B., Gleason, A.C, Wright, E.W., & Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, 106, 623-642.

Obhi, S., & Haggard, P. (2004). Free will and free won't. Motor activity in the brain precedes our awareness of the intention to move, so how is it that we perceive control? *American Scientist, 92*(4), 358-365.

Pockett, S. (2007). The Concept of Free Will: Philosophy, Neuroscience and the Law. *Behavioral sciences & the law*, *25*(2), 281-293.

Roskies, A. (2006). Neuroscientific challenges to free will and responsibility. *Trends in Cognitive Sciences, 10*(9), 419-423.

Stier, M. (2011). Hirnforschung pro Willensfreiheit. Was ist dran? *Nervenheilkunde, 30*, 987-991.

Trevena, J.A., Miller, J. (2002). Cortical Movement Prepraration before and after a Conscious Decision to Move. *Consciousness and Cognition,* 11, 162–190. doi:10.1006/ccog.2002.0548 .

Vohns, K.D., & Schooler, J.W. (2008). The Value of Believing in Free Will. *Psychological Science , 19 (1), 49-54.* doi: 10.1111/j.1467-9280.2008.02045.x

Wegner, D. M. (2004). Précis of the illusion of conscious will. *Behavioral and Brain Sciences, 27*(05), 649-659.