

How Choice Blindness Vindicates Wholeheartedness

ASGER KIRKEBY-HINRUP

Department of Philosophy, University of Lund
Helgonavägen 3, 221 00 Lund, Sweden
asger.kirkeby-hinrup@fil.lu.se

RECEIVED: 12-06-2014 • ACCEPTED: 17-02-2015

ABSTRACT: Recently the account of free will proposed by Harry Frankfurt has come under attack. It has been argued that Frankfurt's notion of wholeheartedness is in conflict with prevalent intuitions about free will and should be abandoned. I will argue that empirical data from choice blindness experiments can vindicate Frankfurt's notion of wholeheartedness. The choice blindness phenomenon exposes that individuals fail to track their own decisions and readily take ownership of, and confabulate reasons for, decisions they did not make. Traditionally this has been taken to be problem for the notion of free will. I argue that Frankfurt's account does not face this problem. Instead, choice blindness can be fruitfully applied to it, and vice versa. Frankfurt's notion of wholeheartedness, I suggest, delineates the range of the choice blindness effect. This makes wholeheartedness a useful meta-theoretical concept for choice blindness research. I conclude that, *pace* the recent criticism, wholeheartedness is a useful notion and should not be abandoned.

KEYWORDS: Choice blindness – decisions – free will – Harry Frankfurt – wholeheartedness.

1. Introduction

Elsewhere (see Kirkeby-Hinrup 2014) I have objected to Harry Frankfurt's account of free will (see Frankfurt 1971; 1988). The objection is that by grounding free will in the notion of wholeheartedness, Frankfurt allegedly renders the notion of free will sparse, practical and occasional. It

becomes sparse because wholehearted identification with a choice rarely obtains in most everyday deliberative situations. It becomes practical because free will essentially depends on particular actions in a context. It becomes occasional because on this view free will only obtains on specific occasions, rather than being something a person can have across situations. I argue that a conception of free will that is sparse, practical and occasional is in conflict with our intuitions: Intuitively we consider free will as possessed by individuals across time. Therefore Frankfurt's notion of wholeheartedness is inadequate. In this article, I will take another perspective and instead of arguing from intuition I take my starting-point in experimental data. I will argue that experiments on choice blindness (cf. Hall – Johansson – Strandberg 2012; Hall – Johansson – Tärning – Sikström – Deutgen 2010; Hall et al. 2013; Johansson – Hall – Sikström – Olsson 2005; Johansson – Hall – Sikström – Tärning – Lind 2006) provides indirect support for Frankfurt's notion of wholeheartedness. In the next section, I will provide a brief introduction to the account of free will developed by Harry Frankfurt. Section three is an introduction to the phenomenon of choice blindness. In section four, I apply the notion of wholeheartedness to the choice blindness phenomenon. This will show that, not only are the two compatible, but from a theoretical point of view they complement each other. In the conclusion, I clarify why my previous critique of the notion of wholeheartedness is less serious than it seemed. Then I explain why the phenomenon of choice blindness is not a threat to accounts of free will grounded in the notion of wholeheartedness.

2. Wholeheartedness and free will

According to Frankfurt, free will is tied to higher-order desires (cf. Frankfurt 1971; 1988). Higher-order desires are desires about (lower-order) desires. Higher-order desires are desires about those desires the individual wishes to be effective. Frankfurt's theory addresses the distinction between simply desiring something and wanting to so desire it. According to Frankfurt, when an individual identifies herself with one of her higher-order desires, such identification will be followed by a higher-order volition. The effect of the higher-order volition is that the individual wishes her particu-

lar desire (the target of her higher-order desire) to be her *effective* desire. An effective desire is a desire that moves her all the way to action. For instance, I may occasionally have a desire for excessive amounts of ice cream. But, for reasons pertaining to health, I rarely wish to act on this desire. Thus, my higher-order desires do not endorse the desire for ice-cream. Consequently, I do not wish this desire to move me to action. Conversely, I have a desire to work out and stay in shape. While this desire may be weak, it is endorsed by my higher-order desires. I wish this desire to be effective, i.e., I wish that I act upon this particular desire. According to Frankfurt, the will is free when the individual acts in accordance with her higher-order volitions.

One objection that is usually leveraged against philosophical accounts invoking higher-order notions is the threat of regress. For Frankfurt's account, the threat of regress pertains to higher-order desires. Is it not possible that a second-order desire could be in conflict with a desire on a level above itself (e.g. a third-order desire)? Why should we not say that if this is the case, then following the higher-order volition, spawned by the third-order desire, is what is necessary for free will? Similarly should we then not say that the fourth-order desires may trump the third-order desires, and so on *ad infinitum*? The problem is that we can always posit a desire of a higher-order. This objection trades on the intuition that the higher the order of a desire, the closer it is to what an individual *really wants*. If such regress is allowed, Frankfurt's account can never get off the ground. The problem is that there is no support for a claim that any particular desire is the relevant one, whose corresponding higher-order volition should ground free will.

To solve the regress problem an account is needed of how an individual comes to identify herself in the right manner with a particular higher-order desire. Frankfurt answers by introducing his notion of wholeheartedness. Frankfurt says:

When a person identifies himself *decisively* with one of his first-order desires, this commitment "resounds" throughout the potentially endless array of higher orders. [...] The decisiveness of the commitment he has made means that he has decided that no further question about his second-order volition, at any higher order, remains to be asked. (Frankfurt 1971, 16 italics from original)

What Frankfurt claims is that a wholehearted identification with a higher-order desire solves the problem. When a higher-order desire resounds throughout the system any question of further higher-order desires is unnecessary. Once the individual has wholeheartedly identified herself with a given desire, she cannot help but to want this desire to be effective. So only when a wholehearted identification with a particular desire occurs is a relevant higher-order volition formed.

To summarize, there are two requisites for free will on Frankfurt's account. The first requisite is that an individual identifies wholeheartedly with some desire, thus forming a higher-order volition. The second requisite is that the eventual decision and action of the individual are in accordance with the higher-order volition. According to Frankfurt, only when both requisites are satisfied, the individual has free will.

3. Choice blindness

Choice blindness experiments expose that people have difficulties keeping track of their decisions. In choice blindness paradigms the individual makes a choice. After making the choice the individual is asked why she preferred the chosen option over the alternative(s). However, in the experimental manipulations the individual is presented with an alternative she in fact did not choose, *as if* she had in fact chosen it. The choice blindness effect is that subjects rarely detect this manipulation. Instead, subjects will confabulate reasons for preferring the option they did not in fact choose. For instance, in one of the early choice blindness studies (see Johansson et al. 2005), subjects are presented with two pictures of individuals of the opposite sex. The subjects are instructed to point to the picture of the individual they find most attractive. After the choice is made (and the subject has pointed), the pictures are placed face down on the table. In the baseline condition, the chosen picture is then picked back up and shown to the subject, while the other picture is still face down on the table. The subject is then asked to describe why she preferred the individual she pointed to, and proceeds to provide an introspective account of her reasons. The experimental condition proceeds in exactly the same way as the baseline condition with the exception that the picture that is picked back up is actually the picture the individual did

not choose.¹ In the experimental condition the subject will normally confabulate reasons for preferring the individual in the picture she is shown.

Since this early study, the choice blindness effect has been demonstrated in a wide range of domains. The various domains in which choice blindness has been demonstrated include the political, moral, aesthetic, gustatory, and olfactory (cf. Hall et al. 2012; Hall et al. 2010; Hall et al. 2013; Johansson et al. 2005; Johansson et al. 2006). Thus, it appears we are blind to the outcomes of a wide range of decisions we normally, and intuitively, take to be important to us. The fact that the choice blindness effect has been demonstrated in different sensory modalities, and across important social domains, such as politics and morals, underscores the pervasiveness of the phenomenon.

Now, it can be argued that choice blindness presents a problem for free will. This is the argument: Given that the choice blindness paradigm has shown that in several domains we are blind to the outcome of our decisions and furthermore that the reasons we provide for those decisions are generated *post hoc*, this may very well be the case with every decision we make. On this interpretation of the outcome of the choice blindness experiments, choice blindness threatens traditional notions of free will because reasons and decisions become epiphenomenal, and thus are outside of conscious control. If the outcome of our choices can be replaced without us noticing, and we readily take ownership of decisions we did not make, this might undermine popular conceptions of free will. Specifically it might undermine the claim that the reasons and deliberations we consciously experience before making a choice have an impact on the choices we end up making. The fact that we experience that our deliberations and decisions matter and are efficacious in our ordinary lives merely may be an illusion.² This suggestion in itself is nothing new (see, e.g. Dennett 1984; 2004; Wegner –

¹ Due to some sleight of hand and the setup of the scenario it appears to the subject as if the experimenter is picking up the correct picture.

² This experience is neatly demonstrated within the choice blindness paradigms as well. When individuals who have just participated in a choice blindness experiment are suggested that perhaps their choices had been switched (which they indeed had), many deny this as impossible and something they would certainly have noticed. This overestimation of own introspective competence has been called *Choice Blindness Blindness* (see Johansson et al. 2005; 2006). Given the choice blindness phenomenon such a naïve faith in one's own capacity to keep track of decisions is clearly mistaken.

Wheatley 1999; Wegner 2002; 2003), but the choice blindness phenomenon *prima facie* provides such a view with more traction. Many accounts (e.g. those of Dennett and Wegner) argue that the experience of free will does not correspond to (let alone entail) any feature of human decision making that corresponds to the experience. These accounts suggest that the notion of free will is either in need of substantial revision (in order to remove the traditional connotations to anything substantively free), or must be abandoned completely.³ The kind of empirical data that forms the basis of such accounts is especially problematic for theories of free will that rely heavily on explicit cognitive processes such as deliberation and introspection.

One option for the proponents of free will is to propose alternative interpretations of the empirical data. In brief, they may suggest that the fact the individuals cannot remember their reasons, entails neither that they did not have any, nor that the ones they did have were inefficacious with respect to their decision. While this defense may be tenable I will not pursue it in the present context. Rather, the objective here is to show that accounts of free will which do not place an emphasis on conscious deliberation, such as Frankfurt's, can sidestep objections based on such empirical data. It will be argued that such accounts can claim that the individual becomes aware of which of her desires are endorsed by higher-order desires by an automatic process. Only the result of the process will be known to the subject. Arguably it is not even necessary that the individual explicitly experiences 'wholeheartedness' consciously. An (unconscious) occurrence of a wholehearted identification would be sufficient.⁴ Furthermore, whether

³ How such revision would influence the view of humans as *free agents* is a separate question. Interestingly, the experience of agency face problems that are similar to the ones choice blindness pose for the experience of free will (see, e.g., Bayne – Levy 2006; Moore – Wegner – Haggard 2009).

⁴ Note that this does not entail that it is futile for the individual to attempt to determine which of her desires she wants to be effective. Wholeheartedness reasonable should ordinarily reveal itself upon inspection. It might be tempting to think that the individual must be conscious of the wholehearted identification in order for this to matter for the decision. However, there are two reasons to hesitate in making this claim. The first reason is that the way Frankfurt explicates his theory, whether the wholehearted identification is conscious, does not seem to matter to whether the individual can exhibit free will. The second is that invoking consciousness of whole-

this process is fully determined is inconsequential to Frankfurt's account of free will. Moreover, there is no requirement that the process must be infallible. That the process may occasionally falter and the individual mistakenly believe that she wants a given desire to be effective is unimportant. This would merely be an instance of those occasions where the individual does not exhibit free will. Because such occasions are allowed on Frankfurt's account, they do not pose a problem.

4. Wholeheartedness and choice blindness

One question about choice blindness that looms large in the background is what kinds of choices can be successfully manipulated. Are there choices the experimenters cannot make subjects believe themselves to have made, and subsequently will not attempt to justify? Surely, intuition suggests, it is impossible to switch a bride at the altar thus tricking the subject into marrying the wrong individual (this intuition is shared by the choice blindness experimenters; see, e.g., Hall et al. 2010). Even more absurd is the idea that the tricked individual would subsequently take ownership of, and justify, the decision to marry the un-intended spouse. If we agree that this is absurd, as I think we should, it follows that there are limits to the choice blindness effect. Simply put, we cannot be tricked to believe we made all choices we are presented with as our own. Can we say anything about these limits?

To investigate the limits of choice blindness is to investigate what characterizes the choices that are immune to the choice blindness manipulation. From Frankfurt's perspective, making his position our own, we can answer this question by deploying the notions of wholeheartedness and higher-order volitions. Because higher-order volitions are about those desires the individual wants to be effective, it seems reasonable to expect that they also track the desires they are about in order to determine if those desires actually *are* effective. How might one argue for the view that such tracking occurs?

heartedness as a necessary aspect, pulls the theory in the direction of the cognitivist accounts of free will, and makes it susceptible to objections from the empirical data that haunt these.

One possibility is the following. Because wholeheartedness, according to Frankfurt, resounds throughout the system, it appears that it cannot be fleeting. To avoid allegations that the notion of wholeheartedness is not robust enough, it seems important for the theory to show how wholeheartedness may persist over time. Now consider the core idea Frankfurt espouses: a higher-order volition is generated when the individual wholeheartedly *identifies* herself with a decision. This idea suggests that there is a link between wholeheartedness and personal identity, in the specific sense of the identity over time provided by the characteristics through which an individual defines herself as herself. I maintain that this link provides the basis and motivation for wholeheartedness to persist over time. For instance, a sincere vegan might take central parts of her identity to consist in particular views on food and related issues.

One might object here that some individuals appear to constantly re-define themselves, while sincerely professing that “this is the new me”. If this is the case, the objection goes, the kind of personal identity referred to above cannot provide the basis for wholeheartedness to persist over time. However, even in such extreme cases, wholehearted commitment to decisions first, does not change on the timescale involved in choice blindness experiments, and second typically occurs in the form of an explicit decision to change one’s mind as opposed to how people change their opinion in the change blindness cases. Furthermore, there is nothing in the notion of wholeheartedness that entails that it can never change. People do change over time. There are beliefs and activities I wholeheartedly endorsed at the age of sixteen I no longer care about. At the very least, the link between wholeheartedness and personal identity provides sufficient reason to believe that individuals can be wholeheartedly committed to the same views or decisions for the entire time a choice blindness experiment runs.

Having made the case that wholeheartedness is geared to persist over time we may now address the question of whether wholeheartedness tracks the outcome of decisions. One way to show that this is true would be to perform choice blindness experiments on subjects who strongly identify with particular views. The idea is that with respect to those views such a subject would be immune to the choice blindness manipulation exactly because wholeheartedness is tied to personal identity. For instance, it seems likely that a professional politician located on the far left of the political

spectrum would immediately object if she were presented with a choice where she supposedly had come out in favor of tougher immigration laws. While it is true that the choice blindness experiments pertaining to political views (i.e. Hall et al., 2013) showed that people on the street were susceptible to the choice blindness effect in this domain, it is unlikely that professional politicians are susceptible to the same extent. This is unlikely exactly because these views normally form a significant part of the how politicians conceive of themselves.

To sum up, it seems that the link between wholeheartedness and personal identity supports the idea that in some cases subjects track their desires and as a consequence the outcome of their decisions. These cases are such that the subject wholeheartedly endorses a particular view or course of action. However, so far choice blindness experiments have not been carried out to test this. To strengthen the case, we turn to another piece of evidence from everyday life that indicates that people tracking their desires and also the outcomes of their decisions.

In our everyday observations of ourselves we often realize when we fail to do what we actually wanted to do. How might this be if we were not keeping track of what we wanted to do? One might counter that such realizations are based on retroactively constructed decisions. However this does not appear to make sense. For instance, taking an evolutionary perspective, what selective pressure would result in a system that retroactively confabulated decisions it failed to follow through on? What would the benefit be? Conversely, it makes sense to track the outcomes of decisions, e.g. for purposes of error detection, learning, and behavioral optimization.

That people occasionally fail to do what they set their mind to is central to Frankfurt's account. The distinction between occasions where we succeed and where we fail to do what we wanted to delineates the occasions where we act on a desire that is not endorsed by a higher-order volition and the occasions where we succeed in acting on the desires we want to be effective. Imagine that I have decided that I really want to work out. To my dismay, I find myself having bought excessive amounts of ice-cream instead. The dismay is an indication that tracking did occur and that another desire was realized than the one that the higher-order desire concerned. I am dismayed because I failed to make the desire my higher-order volition was about (working out) my effective desire. The desire that carried me all

the way to action (eating ice-cream) was not one endorsed by a higher-order desire; it was not one of my concerns.

In this way wholeheartedness demarcates the limit for the choice blindness effect. When I wholeheartedly identify myself with a desire, the higher-order volition tracks or locks on to the desire, and indirectly monitors the outcome of my action. This thwarts the choice blindness effect. We are now in a position to explain the choice blindness effect as occurring whenever decisions are not wholehearted. Because higher-order volitions track their desires, when the outcome concerns another desire, choice blindness occurs. Furthermore, only wholeheartedness generates higher-order volitions, which means that when wholeheartedness is absent there is no tracking going-on and no desire that constitutes a concern for the subject. Consequently, the manipulation of the subject's choice is successful (i.e. not detected) and choice blindness occurs.

5. Conclusion

One might think that intuitively, free will is relevant whenever we deliberate. However, on Frankfurt's account, only the decisions endorsed by a resounding wholeheartedness can be instances of free will. Since the experience of a resounding wholehearted commitment to a given desire or decision is not something that occurs regularly, this means that the majority of everyday decisions are made without free will. While the intuition mentioned above may be one that many agree with, this is not enough to reject Frankfurt's account. Indeed, the fact that the notion of wholeheartedness meshes very well with – and can be fruitfully applied to – empirical data from choice blindness, suggests the intuition might be wrong. If the empirical data show that there is no free will in many trivial choices and across domains, then this suggests the intuition may be wrong. If the data indicate that the intuition may be wrong, then the intuition seems ill suited as the foundation for criticism of Frankfurt's account. On the other hand, seeing the notion of wholeheartedness in light of the phenomenon of choice blindness vindicates the idea that free will might indeed be sparse, practical and occasional (*pace* Kirkeby-Hinrup 2014). Conversely, seeing choice blindness in light of the notion of wholeheartedness suggests that the former not necessarily is a threat to free will. More specifically, it is not

a threat to free will of the kind proposed by Harry Frankfurt. Moreover, while it is likely to be difficult to operationalize empirically, wholeheartedness provides a useful meta-theoretical concept to delineate the limits of choice blindness.

References

- BAYNE, T. – LEVY, N. (2006): The Feeling of Doing: Deconstructing the Phenomenology of Agency. In: Sebanz, N. – Prinz, W. (eds.): *Disorders of Volition*. Cambridge (Mass.): Cambridge University Press.
- DENNETT, D.C. (1984): *Elbow Room: The Varieties of Free Will Worth Wanting*. The MIT Press.
- DENNETT, D.C. (2004): *Freedom Evolves*. Penguin.
- FRANKFURT, H. (1971): Freedom of the Will and the Concept of a Person. *The Journal of Philosophy* 68, No. 1, 5-20.
- FRANKFURT, H. (1988): Identification and Wholeheartedness. In: *The Importance of What We Care About*. Cambridge University Press, 159-176.
- HALL, L. – JOHANSSON, P. – STRANDBERG, T. (2012): Lifting the Veil of Morality: Choice Blindness and Attitude Reversals on a Self-transforming Survey. *PLoS one* 7, No. 9, e45457.
- HALL, L. – JOHANSSON, P. – TÄRNING, B. – SIKSTRÖM, S. – DEUTGEN, T. (2010): Magic at the Marketplace: Choice Blindness for the Taste of Jam and the Smell of Tea. *Cognition* 117, No. 1, 54-61.
- HALL, L. – STRANDBERG, T. – PÄRNAMETS, P. – LIND, A. – TÄRNING, B. – JOHANSSON, P. (2013): How the Polls Can Be Both Spot on and Dead Wrong: Using Choice Blindness to Shift Political Attitudes and Voter Intentions. *PLoS one* 8, No. 4, e60554.
- JOHANSSON, P. – HALL, L. – SIKSTRÖM, S. – OLSSON, A. (2005): Failure to Detect Mismatches between Intention and Outcome in a Simple Decision Task. *Science* 310, No. 5745, 116-119.
- JOHANSSON, P. – HALL, L. – SIKSTRÖM, S. – TÄRNING, B. – LIND, A. (2006): How Something Can Be Said about Telling More than We Can Know: On Choice Blindness and Introspection. *Consciousness and Cognition* 15, No. 4, 673-692.
- KIRKEBY-HINRUP, A. (2014): How to Get Free Will from Positive Reinforcement. *SATS, Northern European Journal of Philosophy* 15, No. 1, 20-38.
- MOORE, J.W. – WEGNER, D.M. – HAGGARD, P. (2009): Modulating the Sense of Agency with External Cues. *Consciousness and Cognition* 18, No. 4, 1056-1064, available at: <http://dx.doi.org/10.1016/j.concog.2009.05.004>
- WEGNER, D.M. – WHEATLEY, T. (1999): Apparent Mental Causation: Sources of the Experience of Will. *American Psychologist* 54, No. 7, 480.
- WEGNER, D.M. (2002): *The Illusion of Conscious Will*. The MIT Press.

WEGNER, D.M. (2003): The Mind's Best Trick: How We Experience Conscious Will. *Trends in Cognitive Sciences* 7, No. 2, 65-69, doi: 10.1016/s1364-6613(03)00002-0.