

A Mixed Judgment Standard for Surrogate Decision Making

Forthcoming in *Journal of Medicine and Philosophy*

ABSTRACT: The Substituted Judgment Standard (SJS) for surrogate decision-making dictates that a surrogate, when making medical decisions on behalf of an incapacitated patient, ought to make the decision that the patient would have made if the patient had decisional capacity. Despite its intuitive appeal, however, SJS has been the target of a variety of criticisms. Most objections to SJS appeal to epistemic difficulties involved in determining what a patient would have decided in a given case. In this paper, I offer an alternative standard for surrogate decision-making that avoids these difficulties. I then offer an account of its theoretical underpinnings which shows that it preserves the central moral justification for SJS, namely, respect for patient autonomy.

§1. Introduction

When a patient is determined to lack decision-making capacity, standard practice is for a surrogate to be appointed and charged with making medical decisions on behalf of the patient. Traditionally, three standards have been used to guide surrogates in deciding on behalf of patients. First, a surrogate may decide on the basis of the patient's expressed wishes regarding her future medical care as they appear in a formal advance directive, typically in the form of a living will. In cases where a living will is available it is taken to be authoritative since, so the claim goes, honoring the expressed wishes of the patient is the best means of respecting the patient's autonomy. In cases where no living will is available, the surrogate is tasked with making the decision that the patient *would have made* were the patient capacitated. This is known as the Substituted Judgment Standard (SJS). Finally, there may be cases in which a surrogate must decide according to the Best

Interest Standard. In such cases the surrogate is required to make the choice that objectively advances the interests of the patient. In practice these principles are often seen as ordered in the sense that one should only be used to guide surrogate decisions if the prior principle fails. (Brock 2004) So, it is generally taken to be the case that surrogates should defer to a patient's advance directive, then, if no such directive exists, the surrogate should apply the Substituted Judgment Standard, and, finally, when this is impossible the decision should be based on the best interests of the patient.

Given that most patients do not have advance directives (Brock 2004), much of the controversy over surrogate decision-making centers on issues in interpreting and applying SJS. Despite several worries about its application (which I discuss at length in what follows), SJS enjoys widespread support for one central reason, namely, that it is seen as respecting the autonomy of the patient by attempting to extend the principle of informed consent beyond the point of incapacity. In this paper, I attempt to offer an alternative to SJS, which I call the Mixed Judgment Standard, that both retains the central moral justification of SJS while avoiding the central objections to it. I begin, in §2, by presenting in further detail a set of important problems for SJS relating to the epistemic limitations of its application. Then, in §3, I present the alternative Mixed Judgment Standard and argue that it is not subject to these same objections and that it respects the autonomy of the patient by ensuring that decisions made on behalf of the patient accord with the patient's values.

§2. Objections to the Substituted Judgment Standard

Many have objected to SJS on the grounds that it presents proxies with an impossible task. We can never truly *know* how an incapacitated patient would have decided, and, so, SJS cannot really extend the autonomy of the patient in the way that its proponents claim. For example, Welie

(2001) has argued that SJS requires that we assume that a patient's past preferences and values remain unchanged and that we are not entitled to this assumption. Welie also raises the objection that since SJS requires that we consult indirect information as a way of gathering evidence of a patient's current values and preferences, it is problematic on this ground as well since the reliability of such information for determining the patient's subjective values and preferences is questionable. We can construct two different versions of this argument based upon what Welie refers to as the two moral foundations for SJS.

The stronger of these two arguments hangs on the claim that the moral basis for deferring to what the patient would have decided is that third parties simply cannot have any independent knowledge of a patient's preferences, values, or interests and, therefore, if we are to do what is best for the patient, then we must act according to the patient's expressed wishes. On this view, individuals are fundamentally "moral strangers." If this is true, then according to Welie, given this pervasive subjectivity, only information that comes directly from the patient can be taken as legitimate for the purposes of surrogate decision making. Thus, he writes,

While one may [if one adopts the moral strangers hypothesis] take into account both written documents by the patient and corroborated verbatim relays of a patient's oral communications, no interpreted or – even worse – intuitive assessments by family, friends and other third persons of a patient's personal values and private interests are admissible. (178)

Insofar as SJS requires the use of such information, then, it lacks the moral justification that its proponents claim.

A full rebuttal of the "moral strangers" view is beyond the scope of this paper. I will assume for the sake of argument, that individuals can have some degree of independent epistemic access to one another's values, preferences, etc.¹ (a point that I return to below). However, even assuming

¹ Indeed, Welie denies the moral strangers approach himself in favor of a view that allows for a degree of intersubjectivity between persons.

this is true, a weaker version of Welie’s objection is still in the offing. If we assume that the moral justification for SJS lies primarily in the fact that we have an obligation to defer to the patient’s autonomous decisions not because doing so is the only means of advancing her interests but merely because they are *her decisions* – this is the other moral foundation identified by Welie – then we may still be left with an epistemic worry about SJS. Namely, if SJS derives its moral justification from the fact that the decision made is *the patient’s* decision, then that justification will be weakened in proportion to the degree of uncertainty that exists as to what the patient would have decided. Moreover, insofar as indirect information about the patient’s wishes is uncertain it will subsequently weaken the moral justification for SJS. Therefore, the epistemic worries about SJS persist.²

It is precisely this decisional uncertainty that forms the basis for a powerful epistemic objection presented by Broström, Johansson, and Nielsen (2007) which puts pressure on the way that SJS is typically formulated. Their central claim is that the counterfactual conditional that SJS requires surrogates to consider is underdetermined and, so, not of any real use in decision making. Broström et al note that all formulations of SJS rely on something like the following counterfactual: *if the patient had capacity, what would she decide?* They then claim that in order to answer this question, we need more information regarding the decision conditions under which

² Edward Wierenga (1983) has offered a somewhat different objection to SJS. He argues that the standard asks us to determine what the patient would decide if she had decisional capacity, but, in many cases (or, perhaps, the majority of cases), the patient would not make the decision that the surrogate recommends because the reason the decision is being made is precisely *because* the patient lacks capacity. We can leave out the details of Wierenga’s argument here, but he argues against the counterfactual conditional in SJS by appealing to a possible worlds semantics. Deborah Barnbaum has argued, plausibly to my mind, against Wierenga’s view in her (1999). There she attributes to Wierenga the following characterization of the counterfactual: “If x’s proxy consent to T on behalf of y is valid, then in the closest possible world in which y is able to offer consent, y consents to T.” (168) On this characterization of the conditional, Wierenga’s conclusion follows. However, according the Barnbaum this characterization is incorrect. Rather, the relevant conditional is the following: “If x’s proxy consent to T in the actual world on behalf of y is valid, then in the closest possible world ... in which y is able to offer consent, y consents to the performance of T in the world in which y has offered consent.” (170) On this reading, she argues, Wierenga’s conclusion does not follow.

the surrogate is supposed to imagine the patient deciding, and they offer several candidate conditions. The options that they offer for providing content to the antecedent of the counterfactual conditional fall into two broad classes. In the first class are conditions that actually occurred at some point in the patient's life, and in the second are conditions that are hypothetical.

In this first class there are several options. We might, for example, be led to make the decision that the patient would have made the last time that she had decisional capacity. On this reading, we would imagine the patient just before she lost capacity (or met some threshold for incapacity) and ask ourselves what she would have decided at that time. However, this is problematic since many patients lose capacity slowly over time, and, as a result, their most recent decision-making capacity may be on the whole worse, and, therefore, less autonomous, than it was at some earlier time in the patient's life. So, perhaps we should instead make the decision that the patient would have made under more favorable conditions, say, before she began to lose decisional capacity at all. Or, instead, perhaps we should make the decision that the patient would have made when her decision-making capacities were at their peak. The problem then becomes that these conditions are likely to yield very different decisions since the decisions made at these points in the patient's life would be influenced by factors external to the patient herself, and we have no non-arbitrary reason for preferring one over the other. So, demanding that the surrogate decide as the patient would have decided *at a given time in the past*, will not, so they argue, yield a determinate decision.

Given the problems with this first class of conditions, we might prefer to fill in the content of the antecedent by appealing to conditions that did not actually obtain at a given time. So, rather than choosing a time in the patient's life perhaps we should make the decision that is most characteristic of the patient's decision-making as it was throughout her life, or for some extended

portion of it. One problem with this way of deciding, however, is that in certain cases the characteristic decision conditions of a person might be unfavorable. Broström et al use a case of a patient who was mildly depressed throughout most of his life to make this point, saying, “this state of mild depression would probably result in an apathetic and indifferent attitude towards having to make up one’s mind at all on an issue like this. Indecisiveness, rather than a definite choice, would be the result.” (273) So, not only might it be the case that this option is indeterminate, it will also likely lead us to an altogether different decision than the first class of options would lead. Broström et al then propose that perhaps what we should really be doing is either constructing a hypothetical version of the patient under mixed conditions that are sufficiently recent, sufficiently favorable, and sufficiently characteristic (in other words, under some combination of past conditions that occurred at distinct times in the patient’s life), or making the decision that the patient would have made under idealized conditions – that is, conditions in which we imagine the patient having ideal decision-making capacities and ideal external circumstances. The problem with both of these options, however, is that both of them ask us to imagine a version of the patient that has never existed, and if the point of SJS is to preserve the autonomy of *the patient*, then it is hard to see how deciding as some imaginary agent would have decided accomplishes this.

What follows from all of this is that without giving some further content to the counterfactual antecedent in SJS there is just no way to answer the question of what the patient would have decided. Moreover, if, as Broström et al argue, there is no theoretical reason to prefer one way of filling out the antecedent over another, then SJS is insufficient to accomplish its goal. In light of this, it might be thought that we ought to abandon SJS in favor of a best interest standard. I think this is the wrong solution, and in the next section, I will propose an alternative standard that avoids the difficulties discussed so far.

§3. The Mixed Judgment Standard

I am sympathetic to the worries posed by Broström et al, and, as a result, I will be concerned in this section with presenting an alternative standard which can overcome these problems while retaining the central moral justification of SJS. The difficulty in filling in the relevant decision conditions in SJS arises from the fact that the apparent task of the surrogate, according to that standard, is to imagine a patient at a particular time prior or to reconstruct some hypothetical version of the patient and to then ask what *that* patient would decide. Problems arise, then, because the *decisional procedure*³ of the patient is opaque. That is, we are unable to reconstruct the decision-making process of the patient. However, attempting to do so, I will argue, is the wrong way to proceed. Rather, what we should be aiming to do is to allow the surrogate and the patient to decide *together*, as it were, what course of treatment the patient should undergo. Therefore, we need what I will call a Mixed Judgment Standard (MJS). The basic proposal is this: when forced to decide on behalf of a patient who lacks capacity, the relevant question for the surrogate is not “what would the patient decide?” Instead, the relevant question is “what would *I* decide if the patient’s values were my own?” That is, the surrogate should take up the evaluative perspective of the patient – she should see the world according to the patient’s values – and use that perspective to supply content to *her own* decisional procedure.^{4,5}

³ At various points I use the terms ‘decisional procedure,’ ‘decision-making process,’ and ‘deliberative mechanism’ interchangeably. What I have in mind here is just the psychological process by which decisions come about for a particular agent.

⁴ Put more colloquially, we might imagine the surrogate saying something like, “If I were her, I would decide x,” where the “if I were her” phrase is taken to mean something like, “if I cared about the things she cares about,” or something similar.

⁵ Here I am making a distinction between the process of deciding and the content of a decision which warrants some clarification. What I have in mind is this: when we make decisions there is a difference between *how* we decide and the *basis* for our decision. We tend to decide on the basis of our preferences, desires, emotional states, and, perhaps most importantly, values. These supply the content for our decision-making, or so I propose. The process for deciding, on the other hand, will involve evaluating the facts of a situation in light of our cares and commitments, identifying reasons for deciding, and weighing reasons against one another in pursuit of some goal. My view, which will be made clearer in what follows is that we can preserve some degree of autonomy by allowing the patient’s values to supply content for surrogate decisions even if we can’t have access to *how* the patient would decide.

In order to defend this standard, I will begin by showing how it avoids the objections leveled against SJS discussed above. I will then argue that, on one influential conception of agency, the mixed judgment standard preserves, as much as is possible, the central moral justification that initially motivated SJS.

§3.1. Advantages of the Mixed Judgment Standard

The central problem for each of the proposed decision conditions discussed by Broström et al is that we are forced to arbitrarily choose a time in the patient's life and to decide as she would have decided at that time or to construct a hypothetical version of the patient and determine how this hypothetical patient would decide. Given that a person's decision-making abilities and tendencies vary widely throughout her life and are affected by her life circumstances, there is little reason to prefer one version of the patient over another. As a result, we lack access to the deliberative mechanism that will allow a decision to be made. This central problem is avoided by the Mixed Judgment Standard in a straightforward way. MJS does not require that a surrogate reconstruct a hypothetical version of the patient or to arbitrarily choose a set of decision conditions. Rather, it merely requires that the surrogate employs her own decision-making faculties under the *actual* conditions in which the decision is being made while using the patient's values as the basis for her own decisions.

Consider the following case as an illustration:

Phil is a retired philosopher. He spent his entire adult life engrossed in research and in teaching his students to engage in the life of the mind. He also cared very deeply for his family and desired their happiness above his own at all times. His life was guided almost entirely by these two values – pursuit of rational inquiry and a deep, abiding love for his family. However, late in life Phil began to suffer from anxiety and occasional panic attacks which had adverse effects on his judgment. He would become fearful and anxious when faced with decisions of even minor importance and would avoid discussing subjects that he found troubling. After a severe stroke, Phil is left permanently unconscious and is being kept alive on a respirator. His oldest daughter is appointed as a surrogate and must decide whether to discontinue Phil's life support.

Applying SJS in this case will leave us with precisely the problem that Broström et al. identify in that we will be forced to choose between the decision that Phil would have made early on in his life and the decision that he likely would have made late in life when he was suffering from anxiety. The former would likely involve removing Phil from the respirator since the values that have guided his life are absent in his present state (and perhaps even subverted if we imagine that the continued expense of his treatment imposes burdens on his family). The latter would likely dictate that we continue use of the respirator since a decision to forego its use would presumably be too daunting for Phil to make given the roles that anxiety and fear have recently played in his decision making. So, as Broström et al. suggest, SJS does not provide sufficient guidance in this case.

However, MJS does provide the needed guidance. Applying MJS, Phil's daughter would need only to ask what *she* would decide if she shared the values of her father. She would need to imagine herself caring about the things that he cared about, and then she would make a decision using her own deliberative mechanism as it is in the actual circumstances. Imagining herself to value family and the life of the mind in the way that her father did, she allows these values to guide her decisional procedure and decides to discontinue the respirator.⁶ The central problem identified by Broström et al., namely, that we lack access to the relevant deliberative mechanism, is avoided because MJS dictates that we substitute the surrogate's deliberative mechanism for the patient's, a mechanism to which the surrogate has full access.

⁶ It has been suggested to me by an anonymous referee that we may have reason to prefer to honor what Phil would likely have decided later in his life if we think of Phil as now *valuing* above all stability and security. As I have imagined the case, it seems unlikely to me that these desires would satisfy the definition of values provided below as it seems that they do not meet the criteria of ultimacy and non-fungibility. However, if we do consider these to reflect new values that Phil has taken on as his life has progressed, then I agree with this conclusion. The important point is that we have access to his values in either case, and, as a result, we have guidance as to how his surrogate should decide on his behalf.

In addition to avoiding the problems raised by Broström et al, MJS also has the virtue of avoiding the problems posed by Welie that I noted above. Recall that Welie argues, in part, that SJS assumes that past preferences of a patient provide insight into the patient's current preferences – that is, that preferences remain unchanged over time – and that the indirect evidence of a patient's preferences is unreliable for the purposes of surrogate decision making. MJS does not fall prey to these same worries. First, according to MJS decisions are not made according to preferences but according to the patient's *values*, and values are much more stable over time than are preferences. While a patient's preferences and values may certainly overlap, the two are not identical by any means. As John Doris puts it, “values are associated with desires that exhibit some degree of strength, duration, ultimacy, and non-fungibility, while playing a determinative-justificatory role in planning.” (2015, 28) This is to say that in order for a desire or preference to be a marker of an agent's values it must be one that persists over time and that is not merely instrumental. One that cannot be replaced without loss and which plays a normative, deliberative role in the patient's actions. Given this characterization, it is safe to say that past values *do* provide evidence of a patient's current values as it is extremely rare for a person to experience a sudden, wholesale change in values.⁷

Even if it is true that past values can give insight into a patient's current values there may be a different objection to MJS in the offing if we grant that values do change over time. Namely, one might wonder if MJS is subject to the same objection that Bröstrom et al. posed to SJS discussed above – what non-arbitrary reason do we have for picking one version of the patient

⁷ It is, of course, not true that this *never* occurs. For example, it might be the case that patients with Alzheimer's dementia experience sudden, fundamental changes in values. Indeed, this is precisely what is at issue in the literature surrounding how to decide on behalf of these patients (see, for example, Dworkin (1986); Jaworska(1999)). These will, no doubt, be hard cases for MJS, and I cannot treat them in any detail here. It will suffice say, for now, that I don't think they will prove to be any *more* difficult than they are for SJS and, so, don't provide reason to prefer SJS.

rather than another? To see why this objection fails when leveled against MJS it will be helpful to more closely examine the reasoning that Bröstrom et al. offer. Their argument seems to go as follows: in order to fill in the antecedent of the SJS conditional we need to specify a set of decision conditions for the patient. The natural starting point, then, is to choose the most recent conditions under which the patient was had decisional capacity. However, Bröstrom et al. argue, these conditions will typically be suboptimal since there was likely some prior set of conditions under which the patient would have been better able to decide. If this is true, then it seems as though we ought to decide as the patient would have under these prior conditions, and once this is granted it will follow that there is no non-arbitrary way of choosing which conditions to prefer. The crux of this argument, then, seems to be the claim that we have good reason to prefer not to choose the patient's most recent *decision conditions*. However, no such reason exists for not preferring the patient's most recent *set of values*. It seems likely to be true that a patient's most recent decision conditions would be suboptimal, but in order for this to be an objection to MJS we must be given an additional argument as to why a patient's most recent set of values is similarly suboptimal. Absent such an argument, MJS is not subject to the same charge of arbitrariness as is SJS.

Second, while we may think that indirect evidence of preferences is unreliable, indirect evidence of values is not. It is fairly easy, in fact, to infer a person's values from her actions and attitudes as values are, in large part, what unifies a person's actions and attitudes over time. It is clearly true that a large part of what motivates others to act and decide in the ways that they do is opaque to us. However, values (understood as having the characteristics specified above) are not nearly so opaque. Given sufficient time to observe and interact with an individual, it strikes me as implausible to think that we could not have a fairly good grasp on what that person values. While it may be true that we cannot know whether any particular action is reflective of an individual's

values or not, given the determinative-justificatory role that values play in planning, it is very likely that an agent's values will be made clear to others over time. In short, actions or attitudes may not inform us of an individual's value set when taken in isolation, but a sufficiently large sample size of actions and attitudes seem very likely to shed light on what a person truly values.⁸

One might wonder whether these advantages of MJS are merely practical or if they actually provide moral reasons to prefer MJS over SJS. In other words, if employing MJS to overcome these practical worries results in decisions that are further away than those that the patient would have made, then the motivation for shifting to MJS would be substantially undermined. Several things may be said in response to this worry. First, it is not clear that MJS *would* result in decisions that differ drastically from the decisions that a patient would actually make, and given its focus on the patient's values, we should not expect this to be the case. This is an empirical claim that would need to be tested, of course, but absent any data there is little reason to think MJS is worse off in this respect. Second, divergence between the surrogate's decision and that of the patient, if it did occur, would not be a problem that is unique to MJS, as SJS has been shown to produce divergence as well. In a review of the empirical data on substituted judgment Shalowitz, Garrett-Mayer, and Wendler found that surrogates fail to make the decision that the patient would have made in nearly one third of cases. When this inaccuracy in SJS is taken in conjunction with the epistemic difficulties outlined above, we have all the motivation we need to shift to a mixed judgment approach. MJS overcomes those epistemic problems and, as I will make clear in the following

⁸ I think that what has been said here gives sufficient reason to believe that the values of others are accessible to us over time. If we take this view, however, we might be left with a different problem. The fact that we can know another's values may lead us to think that we should be more tolerant of paternalistic decision-making by healthcare professionals even in capacitated patients, since we may be able to act on the patient's values without the patient's participation. I think this is incorrect. While we may be able to know the patient's values, this doesn't give us access to *how* the patient would decide on the basis of those values in any ordinary case, and the decisional process is a key component to autonomous decision-making. MJS aims to fill in this decisional gap in patients who lack capacity, but to do so for capacitated patients would be a step too far, even if we know what matters to the patient in question. I'm grateful to an anonymous referee for raising this worry.

section, shifts away from unrealistic attempts to recreate the patient's decision while retaining the central moral justification of SJS.

Before moving on to consider the theoretical justification for MJS, I should pause to note that the view that I am offering may sound similar to a recent view proposed by Phillips and Wendler (2015) which they call the “endorsed life approach.”⁹ Their view, which they take to be an alternative interpretation of SJS rather than an altogether different standard, “understands the SJS as directing surrogates to make decisions based on which option best promotes the life the patient valued for themselves, including the influence the patient wanted to have on the lives of others.” (725) To the extent that Phillips and Wendler take the values of the patient to be the crucial feature of surrogate decision making, I think that they are correct, and clearly my proposed standard shares this feature. However, MJS has a distinct advantage over the endorsed life approach in at least one respect. It will likely be the case that for any surrogate decision there will be competing values at play and that multiple values would have been endorsed by the patient. If the only guidance that the surrogate has is to act according to values that promote the life the patient endorses, then no guidance is given on how to decide between competing, endorsed values. On my proposal, since the surrogate is making the value-based decision “from the inside,” so to speak, having taken up the patient's evaluative perspective as her own, she can decide between competing values by applying her own decisional procedure. Thus, she will have a means of making choices in which competing values are at play.¹⁰ Thus, while a surrogate's using the

⁹ It is worth noting that others have recently proposed alternative standards for surrogate decision-making as well (e.g. Sulmasy & Snyder 2010). I think, unsurprisingly, that MJS is preferable to these as well. However, comparing these alternative standards to MJS would take us too far afield in the present paper, so I will reserve that project for later work.

¹⁰ For further commentary on the endorsed life approach, see Dresser (2015).

endorsed life approach, in my estimation, would respect a patient's autonomy it does not provide actual decision-making guidance, and this gives MJS a considerable practical advantage.

§3.2. Justifying the Mixed Judgment Standard

As I noted at the outset of the paper, the central moral justification for SJS is that it claims to preserve the autonomy of the patient by extending the capacity for informed consent beyond the point of incapacity. Given that the role of the surrogate in MJS is not to ask what the patient would decide but to ask what she, the surrogate, would decide (under the counterfactual conditions laid out above), one might wonder whether MJS retains the moral justification that makes SJS so appealing. In this section, I will argue that while MJS does not perfectly preserve patient autonomy, it preserves what it is that we value about autonomy insofar as the surrogate is able to supplement the agency of the patient in a morally valuable way.

In order to show that this is the case, I will begin by presenting an influential account of agency which was initially proposed by Gary Watson.¹¹ (1975/2004) I cannot do justice to the complexity and nuance of Watson's view here, but providing the basic idea behind it will help to show how MJS might be justified. He begins by noting an important asymmetry between what an agent desires and what an agent values which arises from the fact that, in some cases, the desires that an agent has in no way reflect the agent's values. To use his examples, we can imagine an exhausted mother who suddenly desires to drown her bawling baby in the bathtub or we can imagine a deeply religious person who views his sexual desires as sinful and lacking any value whatever. The possibility of cases such as these, Watson claims, provides the basis for the possibility of unfree action. As he puts the point, "If there are sources of motivation independent

¹¹ The goal of Watson's analysis was to provide a compatibilist account of free action that improves upon hierarchical accounts like the one offered by Harry Frankfurt (1971/1998) rather than to offer an account of autonomous action. It is fairly clear, however, that most actions that are free, in the sense that he is interested in, will also be autonomous. So, the distinction need not trouble us here.

of the agent's values, then it is possible that sometimes he is motivated to do things he does not deem worth doing. This possibility is the basis for the principal problem of free action: a person may be obstructed by his own will." (23) These considerations lead Watson to make a distinction between what he calls an agent's "valuational system" and her "motivational system." This distinction is crucial for the argument of this section, so I will quote Watson at some length.

The valuational system of an agent is that set of considerations which, when combined with his factual beliefs ..., yields judgments of the form: the thing for me to do in these circumstances, all things considered, is *a*. To ascribe free agency to a being presupposes it to be a being that makes judgments of this sort. To be this sort of being, one must assign values to alternative states of affairs, that is, rank them in terms of worth.

The motivational system of an agent is that set of considerations which move him to action. We identify his motivational system by identifying what motivates him. The possibility of unfree action consists in the fact that an agent's valuational system and motivational system may not completely coincide. Those systems harmonize to the extent that what determines the agent's all-things-considered judgments also determines his actions. (25-6, emphasis in original)

So, for Watson, an agent's valuational system is simply the set of judgments that an agent makes regarding which things are worthwhile or "definitive of the good, fulfilling, and defensible life." (25) An agent's motivational system, then, includes all of those desires, attitudes, and considerations which ultimately move the agent to act. Free – or, for our purposes, autonomous – action occurs when these two systems are in accord.

This model of autonomous agency, I think, helps to articulate the moral justification for MJS. If decisions regarding medical treatment are made autonomously, then the considerations that lead one to decide – that is, to act – must be in harmony with (or at least not in conflict with) one's evaluative judgments about the worth of a particular course of treatment (or non-treatment). However, in the case of incapacitated patients what is missing, or impaired, is precisely the ability to act or decide on the basis of *any* considerations or attitudes at all. In other words, patients who lack decisional capacity are deemed incapacitated because of an absence or impairment of the

motivational system. Since, as the above objections to SJS indicate, we cannot recreate the patient's motivational system, MJS dictates that the surrogate's task is to substitute *her own* motivational system for that of the patient. The surrogate takes up the evaluative stance of the patient as her own and brings her motivational system into harmony with that stance. The idea here is that the surrogate, as decision maker, empathically imagines herself as valuing the same things that the patient values. In the same way that an individual acts autonomously when her motivational and valuational systems are aligned, the surrogate respects the autonomy of the patient when her motivational system and the patient's valuational system – which she has empathically adopted – are aligned and this is reflected in her decision.

There are two substantive objections that must be addressed at this point.¹² First, we might wonder why, even if it is possible for the surrogate to substitute her motivational situation for that of the patient, this would be preferable to trying to reconstruct the patient's motivational system. The answer here is that this seems to be the best that we can hope for. I have argued that the patient is not a “black box” when it comes to her values, but it seems to me that she is when it comes to the actual process of making decisions. Broström et al are correct that, in other words, that we are incapable of reconstructing a patient's *decision*. So, MJS aims to get us as close as possible by supplementing the valuational system of the patient with the motivational system of a trusted surrogate.

However, we might still wonder whether this gets us closer to protecting the patient's autonomy insofar as the clinical situation in which surrogate decisions are made is one that is likely to be characterized by grief and stress which would greatly impact the decisional system of the surrogate. This is, indeed, an important concern, but there are two points to make in response.

¹² I'm grateful to two anonymous referees for pressing me here.

First, this is not a mark against MJS *relative to* SJS. The clinical situation is the same in both cases, and it is not clear that it would have a greater effect on the surrogate's ability to decide as she would on the basis of the patient's values than it would on her ability to decide as the patient would have decided. Second, this is not an objection to MJS in principle. Rather, it simply provides reasons to advocate for surrogate support in order to create conditions that are conducive to sound decision-making.

The second substantive objection is this: one may doubt that surrogates are actually capable of meeting the standards of MJS. It may be possible to have epistemic access to another person's values, but we might doubt whether one can actually *take on* those values in any robust way. Insofar as values are a central component of one's identity, we may be dubious as to whether one could substitute another's values for her own. There may be something to this objection, but it is one that, I think, MJS can handle. We need not require that a surrogate renounce her own values and fully takes on those of the patient – to ask this likely would be to ask too much. All that is required for MJS is that the surrogate be able to *identify* with the patient to the point of taking up the patient's perspective to some substantial degree. To ask this of a surrogate is to ask something of which most of us are quite capable. To take on the patient's evaluative perspective in this sense is, fundamentally, to identify empathically with the patient. A robust capacity for empathy is something that most surrogates have and that allows for taking up the value-perspective of others.¹³ Moreover, empathic identification need not require that one loses a felt sense of independence from the target of her empathy.¹⁴

It is important to note that this approach does not, strictly speaking, *preserve* the patient's autonomy since the decision being made involves the considerations of two separate agents. At

¹³ See Shoemaker (2015)

¹⁴ For more on this see Slote (2010, ch. 1)

best we might call this decision “jointly autonomous,” or, perhaps, “*d*yonomous.” Nonetheless, to the extent that the desires, cares, and concerns that lead the surrogate to decide accord with the valuational system of the patient, dyonomy is the closest thing to autonomy that can be hoped for.¹⁵ Importantly, however, this, it seems to me, is sufficient moral justification for MJS. The reason that we care about autonomy as much as we do is precisely that we think that persons have the right to live the life that they deem most valuable. Certainly, it matters to us that we be allowed to act or decide without the interference of others, but this is secondary to the moral significance that we place on living according to what we value. To the extent that MJS preserves *that*, it has all of the justification that it needs, and, given that it solves many of the problems associated with SJS, it is clearly preferable as a standard for surrogate decision making.

§4. Conclusion

In this paper, I have argued for a new standard for surrogate decision making. According to that standard, the Mixed Judgment Standard, rather than answering the question, “what would the patient decide if she were able to make decisions for herself?” the surrogate is directed to answer the question “what would I decide in this situation if the patient’s values were my own?” This standard avoids the epistemic problems that face the traditional substituted judgment standard, and, I have argued, it retains the central moral justification that makes SJS so appealing all while offering more guidance on how to select a surrogate for a given patient. All of this speaks in favor of moving away from SJS in practice and adopting MJS as an alternative.

¹⁵ Interestingly, it may be possible to construct a similar justification on the basis of a hierarchical view of autonomy like that proposed by Gerald Dworkin. On his view, “autonomy is conceived of as a second order capacity of persons to reflect critically on their first order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values.” (1988, 20) On this conception of autonomy, we might say that a surrogate should attempt to take on the patient’s higher-order attitudes and then act on the desires that she would have in light of them. However, I think the Watsonian approach is far more promising, both for the reasons that Watson offers against hierarchical views and because, as I have suggested already, values are far more easily discernible than are preferences or desires, even those of a higher order.

REFERENCES

- Barnbaum, Deborah. "Interpreting Surrogate Consent Using Counterfactuals." *Journal of Applied Philosophy*. Vol. 16, No. 2. (1999): 167-172.
- Brock, Dan W. "Surrogate Decision-Making." In *Encyclopedia of Bioethics, 3rd Edition*, 2483–2486. Edited by Stephen G. Post. New York: Macmillan, 2004.
- Broström, Linus, Mats Johansson, and Morton Klemme Nielsen. "“What the Patient Would Have Decided”: A Fundamental Problem with the Substituted Judgment Standard." *Medicine, Health Care, and Philosophy*, Vol. 10. (2007): 265-278.
- Doris, John M. *Talking to Ourselves: Reflection, Ignorance, and Agency*. Oxford: Oxford University Press, 2015.
- Dresser, Rebecca. "Substituted Judgment in Real Life." *Journal of Medical Ethics*, Vol. 41. (2015): 731-732.
- Dworkin, Gerald. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.
- Dworkin, Ronald. "Autonomy and the Demented Self." *The Millbank Quarterly*, Vol. 64, Supplement No. 2: *Medical Decision Making for the Demented and Dying*. (1986): 4-16.
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person." In *The Importance of What We Care About*, 11-25. Cambridge: Cambridge University Press, 1998.
- Jaworska, Agnieszka. "Respecting the Margins of Agency: Alzheimer’s Patients and the Capacity to Value." *Philosophy and Public Affairs*, Vol. 28, No. 2. (1999): 105-138.

- Phillips, John and David Wendler. "Clarifying Substituted Judgment: the Endorsed Life Approach." *Journal of Medical Ethics*, Vol. 41. (2015): 723-730.
- Shalowitz, David I., Elizabeth Garrett-Mayer, and David Wendler. "The Accuracy of Surrogate Decision Makers: A Systematic Review." *Archives of Internal Medicine*. Vol. 166, No. 5. (2006): 493-497.
- Shoemaker, David. *Responsibility from the Margins*. Oxford: Oxford University Press, 2015.
- Slote, Michael. *Moral Sentimentalism*. New York: Oxford University Press, 2010.
- Sulmasy, Danieal & Lois Snyder. "Substituted Interests and Best Judgments: An Integrated Model of Surrogate Decision Making." *Journal of the American Medical Association*, Vol. 304, No. 7. (2010): 1946-1947.
- Watson, Gary. 1975. "Free Agency." In *Agency and Answerability: Selected Essays*, 13-32. Oxford: Oxford University Press, 2004.
- Welie, Jos V. M. "Living Wills and Substituted Judgments: A Critical Analysis." *Medicine Health Care, and Philosophy*, Vol. 4. (2001): 169-183.
- Wierenga, Edward. "Proxy Consent and Counterfactual Wishes." *Journal of Medicine and Philosophy*. Vol. 8, No. 4. (1983): 405-416.