

In Their Own Best Interest. Is There a Paternalistic Case for Welfare Conditionality?¹

ANDERS MOLANDER*  AND GAUTE TORSVIK**

*Centre for the Study of Professions, Oslo Metropolitan University

**Department of Economics, University of Oslo and Centre for the Study of Professions, Oslo Metropolitan University

Abstract

This paper examines paternalism as a justification for welfare reforms making benefits conditional on participation in activation programs. We clarify different types of what we denote ‘throffer paternalism’ – a paternalism conjoining an offer with a threat – and ask whether there is a good case for any of them. We argue that hard but non-perfectionistic paternalism provides the most promising defense for mandatory activation but conclude that it does not give a convincing justification for this type of welfare policy.

Keywords: welfare conditionality; welfare reform; paternalism; perfectionism; throffer

Introduction

Since the 1990s, there has been a movement across welfare states toward activation policies (Eichhorst et al., 2010). Activation policies come in different forms and shapes, but, despite national differences, a central common feature of these policies is that welfare recipients are obliged to participate in measures that should enhance their prospects to get paid work.² The obligation to participate implies sanctions if the activation requirements are not followed. This means that welfare recipients are presented with a ‘throffer’, to use a term coined by the philosopher Hillel Steiner (Steiner, 1974–75; see also Taylor 1982). Those who do not accept the offer (to participate in a work-related program) are faced with the threat of having their benefits reduced or withdrawn. It is this combination of threat and offer that makes activation policies controversial (Goodin, 1998).

Different types of arguments can be used to defend activation throffers (Molander and Torsvik, 2015). One may argue in terms of economic efficiency (the collective gain of reducing the number of persons on welfare is higher than the costs of running the programs), legitimacy (increased support for the benefit schemes because mandatory activation may help target assistance to those who cannot get a job), or fairness (it is fair to demand an effort from people receiving public income support; those who can work, with or without training, but do not

want to make an effort *de facto* exploit those who contribute to the system of public income support). These arguments are all other-regarding in the sense that they are concerned with how people who are not subject to activation policies are affected. A fourth argument focuses on the targeted welfare recipients themselves and claims that mandatory activation is in their own best interest. In political philosophy this argument goes under the name of ‘paternalism’.

These four arguments constitute a ‘justificatory set’ that political stakeholders can rely on in their construction of ‘justificatory narratives’ (Eriksen and Molander, 2019). Such narratives arrange the arguments, but seldom elaborate on them. This vagueness may be a rhetorical advantage in public political debate. It can hide disagreements and make arguments appear common sense-like. However, to assess arguments from a substantive point of view – whether they are valid or tenable – we need to present them in a more systematic form than the one in which they usually appear. It is through such clarification or rational reconstruction that political philosophy can contribute to a better understanding of what is at stake in political debates (Swift and White, 2008).

In this paper we address the paternalistic argument and dissect the conditions that must be met in order to provide a tenable justification for mandatory activation. Although policy stakeholders, for obvious reasons, steer clear of the term ‘paternalism’, they often refer to welfare recipients’ own good in some combination with other arguments when they are asked to justify conditional benefits (Dwyer, 2020, Eriksen and Molander, 2019, Mestan, 2016). In the social policy literature, Lawrence Mead’s “new paternalism” counts as the canonical defense of sanction-backed behavioral requirements for welfare recipients (Watts and Fitzpatrick, 2018, Mead, 1997 and 2005). However, Mead gives no more detailed justification of the welfare policy he advocates and does not relate in any significant way to the long and extensive discussions of paternalism in legal and political philosophy. The latter also applies to those who denounce mandatory activation policies as “neoliberal paternalism” (Soss et al., 2011).

Despite a vivid philosophical discourse of paternalism, philosophers have so far shown little interest in the case of welfare conditionality (see, for example, Grill and Hanna, 2018). With this paper, we wish to remedy this lack of communication between social policy and the philosophy of paternalism. Our aim is to clarify and evaluate a paternalistic justification for the use of ‘throtters’ in income replacement schemes. Fully aware of the fact that this argument appears together with other arguments, we examine it separately in order to assess its tenability *per se*, as has been done in detail with the reciprocity argument (White, 2008). We differentiate between types of paternalism that can be invoked to justify mandatory activation and ask whether there is a good case for any of them. Our contention is that hard but non-perfectionistic paternalism provides the most promising defense for welfare conditionality but conclude

that it does not give a convincing justification for imposing mandatory activation requirements on welfare recipients.

Throffer paternalism

A threat states an intention to make someone worse off. An offer proposes an option. You can decline an offer and choose to remain in the situation prior to the offer. A ‘throffer’ threatens to make you worse off if the offer is rejected (Wertheimer, 1987: 204 ff.). Mandatory activation programs are throffers because the offer to take part in activities is conjoined with the threat that those who refuse to take part will be sanctioned. The coerciveness of a throffer depends on what options it leaves (Shapiro, 2007: 216 ff.). If someone hold a gun to your head and says, “I’ll pay you 1000 dollars for the job; if you don’t do it, I’ll kill you”, this is clearly coercive. Activation throffers are less coercive. It is possible to decline the offer to participate in the activation, take the sanction and still survive. A civilized welfare state will not let its citizens starve, but those who do not have family or friends to lean on do not have much of a choice when facing an activation throffer.

In general, paternalism is the view that it is permissible for an outside agent to induce a person to do something that, according to the agent, is for this person’s own good. There has been considerable debate on how to define paternalism more precisely, reflecting different views on whether – and if so, under what conditions – paternalism is legitimate.³ Paternalism can be coercive or non-coercive; it can be practiced person to person, or by professionals in their interactions with clients and patients; or by private or public institutions. The kind of paternalism that is relevant in the case of welfare conditionality is so-called legal paternalism: that is, coercive interventions by public institutions (or their representatives) that are authorized by law (Feinberg, 1986: 7–9). Interventions are coercive when threats are employed to force people to act in a certain way. The idea of ‘nudging’ (or ‘choice architecture’) introduced by Thaler and Sunstein does not belong to this kind of paternalism because the interventions are non-coercive and aim to be compatible with freedom of choice (Thaler and Sunstein, 2008). What distinguishes legal paternalism from other forms of legal coercion is that it is based on a negative judgment of the targeted person’s ability to make a prudent, self-regarding decision and that the interventions therefore are justified by reasons claiming that they will protect a person from self-harm and make him or her better off. Paternalism is at odds with the liberal principle, stated by John Stuart Mill, that “/t/he only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant.” (Mill, 1959/1989:13).

The classical liberal rejection of legal paternalism is grounded in the negative rights of citizens – i.e., the right to freedom from interference from other citizens or the state – and the state’s corresponding negative obligations not to interfere within the sphere defined by these rights (Hösle, 1997: 802). Throffer paternalism, as a kind of legal paternalism, has as its basis a positive right – a claim to some good or service – or, from the state’s perspective, a positive obligation to provide a good or service. This separates it from, for example, legal prohibitions on drugs, bans on unhealthy food, or the required use of seat belts or helmets. The targets of welfare conditionality receive benefits the state provides as a right, and activation measures aim at enhancing their prospects of self-sufficiency. Other examples of throffer paternalism include drug addicts who must accept a sobriety treatment to receive economic support and health services, or patients who must participate in an exercise or diet program to receive specific medical treatment. Throffer paternalism rests on the presumption that, if the state provides a positive right, it is also permitted to require a certain self-regarding conduct by its recipients and threaten to withdraw or limit this right in the case of non-compliance. We can thus define throffer paternalism as the authority (a) to enforce welfare recipients’ participation in offered activities by threatening to reduce or withdraw benefits to which they are initially entitled, (b) and to do so because participation in these activities is in their own best interest.

Douglas MacKay argues that a definition of paternalism, which includes interference with liberty (Cf. Dworkin, 1972: 65), does not apply to conditional welfare state programs since they provide goods or services (MacKay, 2019). However, MacKay does not take the throffer mechanism into account. Even if a throffer does not entirely block the option to decline the offer, it makes this choice costly, and sometimes so costly that there is no real choice. What makes welfare conditionality paternalistic is not the throffer per se, but that it is justified by the welfare recipients’ own good. As we mentioned initially, there are also other reasons for making benefits conditional on behavioral requirements, and paternalistic reasons for mandatory activation can be – and most often are – combined with other justifications for such policies.

Legal paternalism, including throffer paternalism, can take different forms. Two often-made distinctions are between perfectionistic and non-perfectionistic paternalism, also denoted ends- and means-related paternalism, and between hard and soft paternalism.⁴ Perfectionistic paternalists believe that it is possible to determine what is good for a person independently of this person’s own standards. Some activities, experiences, relationships and ways of life constitute a good or meaningful life. They are worthy not because they are desired, but because they are worthy of being desired, and it is a legitimate aim of the state to use its power to promote these endeavors and discourage those who detract from the good life (see Raz, 1986; Wall, 1998). This justification for paternalistic interventions is inevitably controversial in a pluralistic society and therefore in

tension with the liberal principle of state neutrality according to which the state should restrict itself to establish a fair framework for cooperation within which citizens can pursue their conceptions of a valuable life (see Rawls, 1993; Quong, 2011). When public power restricts individual liberty with reference to conceptions of the good that are subject to reasonable disagreement citizens are not respected as free and equal persons. However, there are liberals who reject the principle of neutrality and attempt to combine liberalism's commitment to individual freedom with perfectionism by ascribing intrinsic value to personal autonomy (Raz, 1986; Wall, 1998). According to liberal perfectionism, leading a good life means leading an autonomous life. The ideal of autonomy is the vision of "people controlling, to some degree, their own destiny, fashioning it through successive decisions throughout their life." (Raz, 1986: 372–3) For a liberal state, it is permissible to promote this ideal, which can be done through a wide set of measures, including redistributive policies, educational and cultural policy, as well as through interventions preventing individuals from destroying their capacities for an autonomous life. Defenders of liberal neutrality regard liberal perfectionism as too comprehensive. By favoring a particular conception of the good life, the state goes beyond matters of justice (Quong, 2011). Moreover, liberal perfectionism is an unstable position because the ideal of leading an autonomous life also justifies negative liberty and can serve as a justification for anti-paternalism.⁵ It is not clear how liberal perfectionism trades off negative liberty against interventions serving the purpose of enhancing the conditions for autonomy, and what range of paternalistic interventions that is licensed. In fact, liberal perfectionism can advocate quite far-reaching limitations on personal freedom.

Non-perfectionistic paternalism appeals to people's own understanding of what is good for them, and justifies coercive interventions when people make choices that do not further their own ends.⁶ The aim is not to impose a view of the good life but rather "to make them better off, as judged by themselves" (Thaler and Sunstein, 2008:5; Sunstein, 2016: 43–52) or "getting people to where they want to go" (Conly, 2013:33). The range of interventions supported by non-perfectionistic paternalism differs according to the interpretation of what is good for people as per their own understanding. A critical distinction is whether coercion is defensible only if individuals act in a way that diverges from their current preferences, or if coercion is permissible also to align choices with what individuals would have chosen if they were more informed and made better use of their capacity for instrumental reasoning. The range of interventions that is permissible expands as the interpretation of 'better off' moves in the direction of the presumed choices of a 'well-informed', 'farsighted' and 'instrumentally rational' person. When appealing to such hypothetical choices, non-perfectionism faces an epistemic problem. How can we possibly know that our reconstruction of what people would ideally want is adequate? How can we know that the

targets of a paternalistic intervention would consent after sober reflection? (see Sugden, 2019: ch. 4; White, 2011, ch. 5)

Hard paternalism holds the view that it is permissible – or even required – for the state to limit the liberty of persons if it is done to protect them from the self-harming consequences of their actions, even if these actions are fully deliberate and voluntary (Feinberg, 1986: 12). The sole criterion for intervention is ‘harm to self’. Soft paternalism adds another criterion relating to the conditions under which the actions are undertaken. It considers paternalistic interventions to be legitimate only when individuals harm themselves because of ignorance, impaired judgmental ability, systematic reasoning failures, or weakness of will. While hard paternalism trades off potential self-harm against personal liberty, soft paternalism does so only after having balanced potential self-harm against the capacity of choice. The scope of soft paternalistic interventions depends on where the soft paternalist sets the standard for harm to self and draws the threshold below which it is permissible to interfere with potentially self-harming choices. If the threshold for being a rational or competent chooser is set high, many choices can become candidates of such interventions, and soft paternalism is in danger of entering a slippery slope.

This results in four types of paternalism and throffer paternalism:

	Hard paternalism	Soft paternalism
Perfectionism	1	2
Non-perfectionism	3	4

The strongest version of throffer paternalism (1) argues that welfare conditionality is a legitimate way of getting welfare recipients to achieve self-reliance as an end in itself. The weakest version (4) refers only to welfare recipients’ own standards – actual or ideal – and ‘throffering’ is permissible only if non-participation is based on judgmental failures. Version (3) also argues that welfare conditionality just gets people ‘where they want to go’ but does not, in contrast to (4), introduce any judgmental criteria: it is only potential ‘harm to self’ that matters, not under which conditions the choice not to participate is made. Regarding (2), this is hardly a consistent position. Soft paternalism assumes that the targets of welfare conditionality would accept to be ‘throffered’ if they were more rational or had more self-control, while perfectionism holds that what is good for them can be fixed irrespective of their beliefs and desires.

Standard objections to paternalism

There are two standard objections to the paternalistic use of state power. The first is based on the basic liberal principle that the state should guarantee the

greatest possible freedom for each that is compatible with the same freedom for all. Government interventions should protect the equal freedom of individuals and prevent them from causing harm to each other. Their own good is not a public concern, as Mill stated: “Over himself, over his own body and mind, the individual is sovereign.” (Mill, 1859/1989: 13) When public institutions interfere with the self-regarding choices of mature persons, they act wrongfully because they violate a sphere of legitimate individual agency: individuals have a right to pursue their own ideas of the good life in their own way. Of course, individuals may make imprudent choices. However, they should be free to pursue their own good in their own way and not be forced to act otherwise as long as they do not harm others. What is wrong with paternalism is not only that people may experience interference with their choices as insulting, but the moral status of mature persons is also demeaned when they are treated as though they lack the ability to choose for themselves (Shiffrin, 2005). There is, so the argument goes, an intrinsic relationship between freedom of choice and what constitutes our humanity. Therefore, to be deprived of this freedom is to be deprived of one’s human dignity.⁷ This argument from autonomy can be developed into a public reason argument holding that restrictions on individual liberty must be justified on grounds that can be accepted by citizens who recognize their common moral status as free and equal persons (see Quong, 2011).

The other standard objection is epistemic and insists that individuals generally know their own best interests. Paternalism is likely to interfere ‘wrongly, and in the wrong place’ because of information asymmetries between public institutions and their officials, on the one hand, and the targeted persons on the other (Mill, 1859/1989: 83; Sunstein, 2014: 91–102). Rather than promoting people’s welfare, paternalism will most likely make their lives worse. Moreover, when governments interfere with self-regarding choices, they may not only err but also hinder the development of an individual’s ability to choose how to live one’s life. This ability is, as Mill said, “like the muscular powers . . . improved only by being used.” (Mill, 1859/1989: 59)

Of the two objections, the first is the more basic, since it implies that paternalism is wrong even if a paternalizer is right about what benefits other persons. According to this objection, what is wrong with paternalism is “not simply misdirected care or even negligently misdirected care. It is, rather, primarily a failure of respect, a failure to recognize the authority that persons have to demand, within certain limits, that they be allowed to make their own choices for themselves.” (Darwell, 2006: 268)

The second objection makes a problematic general claim about individual choice. We cannot assume that individuals always, or most of the time, know what is in their own best interest. Moreover, if they know, they may suffer from weakness of will. Research in psychology and behavioral economics has shown that people suffer from systematic reasoning failures and lack of self-control

(Thaler and Sunstein, 2008). This fact is the starting point for those who allow for some degree of soft paternalism. They make a distinction between autonomy as freedom of choice, or negative liberty, and autonomy as a capacity for making deliberate choices. Gerald Dworkin summarizes autonomy in the latter sense “as a second order capacity of persons to reflect critically upon 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” (Dworkin, 1991: 20). This capacity for autonomy may fail due to various factors. According to soft paternalism, it is permissible for public institutions to interfere with personal autonomy when individuals’ self-harming choices cannot be considered truly their own because of ignorance, reasoning failures, or *akrasia*, weakness of will. Thus, it is an exception approach (Hanna, 2018b: 146). There is a strong presumption in favor of freedom of choice, but it can be overruled when self-harming choices are made under somehow distorted conditions of judgment. Although soft paternalism is a view that many liberals embrace, and some even dispute whether it is paternalistic at all (Feinberg, 1986: 12–13), it involves this fundamental problem: How can we know that a person’s choices are flawed according to his or her reflective values and not explained by this person – deliberately – making other trade-offs than the paternaliser? Predictions about what a person would have chosen if they were fully informed are fallible, and there is a risk for well-meaning but – according to soft paternalism’s own standards – unjust interventions.

A paternalistic case for welfare conditionality?

Let us now turn to the question of whether welfare conditionality can be justified on paternalistic grounds. As we have pointed out, welfare conditionality belongs to a context of positive rights and obligations, and the paternalist will argue that the state is permitted to require a prudent self-regarding conduct in return for guaranteeing income support as a matter of right. From here, a paternalistic argument can take different paths. We shall first distinguish between perfectionist and non-perfectionist versions of activation paternalism, then proceed to the distinction between hard and soft paternalism and conclude by stating what we take to be the most convincing form of throffer paternalism. However, we will argue that even this position falls short of providing a tenable defense of welfare conditionality.

a. Perfectionism–Non-perfectionism

A perfectionistic argument for welfare conditionality would claim that work-related requirements bring people closer to economic self-reliance, which is a valuable way of living, irrespective of the welfare recipients’ own attitudes and standards. This argument is based on a rather restrictive idea of what

constitutes a good life, taking into consideration all the beings and doings we have reasons to value. To make self-reliance an end in itself, irrespective of the type of jobs a person can realistically get and how leisure time and other activities are valued, is not very convincing. Some would even consider self-sufficiency a crazy criterion for a good life.⁸ Moreover, many wealthy individuals appear to be living good work-free lives.

However, a broader perfectionistic argument can be constructed with reference to the non-instrumental aspects of employment, such as those that Marie Jahoda specified on the background of her life-long research on the social psychology of unemployment. According to her, unemployed people lack time structure, collective purpose, social contact, status, and activity (Jahoda, 1982). However, these aspects are only ‘latent functions’ of having a job, the manifest being earning an income, and as such – as essentially by-products – they cannot be said to constitute the “intrinsic” value of work. A loftier argument would invoke an idea of work as self-realization: that is, as an activity that is constitutive of a true human life.⁹ However, this romantic idea envisions work conditions that are very different from what most of the jobs provide that are within the range for those who are targets of mandatory activation. Rather than justifying conditionality, the idea of self-realization through work would lead to a criticism of what the young Marx called ‘alienated labor’ under capitalism (Marx, 1844/1975).

Perfectionistic justifications of conditionality will conflict with the liberal principle that the state should not enforce any specific conception of the good life but rather provide fair conditions for social co-operation. People may, of course, value self-reliance, adopt a work ethic in their personal life, or subscribe to a vision of self-realization through work. However, if public institutions impose this way of life on citizens, they do not respect them as being free and equal (see Van Parijs and Vanderborght, 2017:99–100, White, 2005: 89–91). Moreover, it is insulting to citizens who are unwillingly dependent on income support, and who try to live their lives as well as possible. For these reasons, it is difficult to imagine a viable *public* justification of welfare conditionality based on a specific work-related conception of the human good.¹⁰

A more viable perfectionist defense of welfare conditionality may refer to the ideal of personal autonomy, as liberal perfectionism does. To be autonomous, individuals must not only have certain mental abilities but also a will to live a self-directed life and an adequate scope of choice. To the extent that self-reliance is better than being dependent on public support regarding the possibility of leading an autonomous life, welfare conditionality can be justified as an autonomy-enhancing policy (Henning, 2009). However, even if one accepts that there is a positive link between paid work and autonomy, the question remains whether this kind of perfectionism is coherent. First, the ideal of an autonomous life is a particular idea of a good life, and to impose it would be

as illiberal as using state power to force people to live according to some other particular conception of the good life. Second, how can the value of autonomy provide reasons for both the restrictions on the use of state power – that is, guarantees for citizens’ negative liberty that leave them free to pursue their own good – and for coercive measures preventing them from making imprudent self-regarding choices?

For a liberal perfectionist, the goal of enhancing autonomy may justify policies for bettering the labor market opportunities of welfare recipients, and make it permissible to tax citizens: that is, to interfere with their liberty, to finance public programs qualifying people for the labor market. However, it is uncertain whether a liberal perfectionist would accept that participation in such program should be mandatory. A liberal perfectionist would probably prefer to give welfare recipients offers, not confront them with thoffers.

A non-perfectionist defense of mandatory activation does not start from a pre-established conception of the good but will appeal to the target’s own standards. The purpose of a thoffer is to help them to achieve their own ends. Welfare recipients who would not want to participate in work-related programs if they were offers rather than thoffers, would presumably act in a way that is inconsistent with their own values. As Mead puts it in his defense of workfare, “In fact, the clients of paternalism commonly do accept the values enforced; especially, they express a desire to work. However, they commonly fail to conform to these values in practice. Paternalism seeks to close that gap.” (Mead, 1997:5). The problem is, according to Mead, a “lack of competence” or of those “qualities that allow a person to get ahead in economic terms” (Mead, 1997:19), and here paternalistic interventions should help them close the gap by acquiring this competence. The challenge of this argument is, of course, the presumption that welfare recipients ‘commonly’ desire to work. However, what if this is not true? What if there are a lot of welfare recipients who want to live on benefits and are voluntarily out of work? (Dunn, 2014) In this case, a paternalistic justification for mandatory activation must either go in a perfectionistic direction and appeal to an ‘objective’ conception of the good, or appeal hypothetically to what welfare recipients would desire if they could weigh all relevant information in an unbiased way. Following the latter non-perfectionistic line, one may argue that reasonable and well-informed welfare recipients – all things considered – prefer to have a job above living on public benefits. When reasonable and well-informed, they would willingly take part in work-related programs that increase their chances of getting a job (White, 2008:148). Based on this reconstruction of what would be an informed and reasoned judgment, those who reject taking part in employability-enhancing measures act against their best interest. The crux here is how to infer that such an idealized hypothetical judgment would be the judgment of real-world welfare recipients in a state of sober reflection. How can we know that they would consent to be ‘thofferred’ even though they

reject the offer? If judgments about best interests of hypothetical 'rational' welfare recipients are not the considered judgments of ordinary welfare recipients, then this allegedly non-perfectionistic defense of conditionality is only a less blatant or straightforward form of perfectionism. Furthermore, even if a second-order or latent preference for employment and self-reliance has been reconstructed, welfare recipients may choose not to participate in activation programs because they believe the costs outweigh the benefits, not because they suffer from weakness of will.

b. Hard and soft paternalism

According to the hard paternalist, welfare conditionality is required because welfare recipients are believed to be harming themselves if they do not participate in work-related programs. The hard paternalist can have both a perfectionist (ends-oriented) and non-perfectionist (means-oriented) account of the value of self-reliance. A perfectionist account is difficult to defend for the reasons mentioned above. However, a non-perfectionist account also has its problems, as we have seen.

For the soft paternalist the legitimacy of mandatory activation depends on the choice capacity of welfare recipients. He or she makes two assumptions. First, there are true gains from participating in activation programs that surpass the costs associated with the loss of personal freedom that comes from forced participation. Second, welfare recipients would not participate if the programs were voluntary because they are confused about their own best interests or suffer from *akrasia*. In the negative judgment of welfare recipients' ability to make the prudent choice, the soft paternalist cannot appeal to an ideal of self-reliance, which would imply a perfectionist stance that is inconsistent with the logic of soft paternalism. A soft paternalist must refer to what would be the choice of a reflective and prudent welfare recipient. However, as noted above, a key question is: How can the soft paternalist know what would be the rational choice of a person who lacks these attributes? Is the hypothetical consent simply the soft paternalist's own judgment about what welfare recipients should agree to? If so, what sets soft paternalism apart from hard paternalism?

c. The best account of throtter paternalism

Jonathan Wolf has argued that activation policies imply perfectionism: for example, by referring to an ethics of work (Wolf, 2004). However, this is not the case. Mandatory activation can be defended with other arguments: for example, with reference to an idea of economic efficiency or based on the conception of justice as reciprocity. Furthermore, nor is it true that a paternalistic justification must take a perfectionistic stand on the human good. As we have seen, both hard and soft paternalism can make the more cautious claim that taking part in a work-related program is something that will improve recipients' welfare

prospects according to their best judgment. This is also the most viable paternalistic defense given the plurality of reasonable ideas of what constitutes a good life. Therefore, if we limit our discussion to the realm of non-perfectionistic paternalism, which variant of paternalism – the hard or the soft – is best suited to defend mandatory activation?

Soft paternalism differs from hard by including choice capacity as a criterion for whether paternalistic interventions are legitimate. It makes exceptions from the Millian harm principle when self-regarding choices are subject to rationality failures and weakness of will. For example, a soft paternalist may, therefore, find it acceptable to force an addict to take part in a weaning program. Does a welfare recipient who declines the offer to take part in an activation program constitute an analogous case? Is it reasonable to assume that their choice capacity is seriously damaged, that their disadvantageous background has made them disoriented and demotivated and therefore unwilling to take part in a program that they will benefit from in terms of employability and future welfare? Hard paternalism does not require considerations of this kind. If the individual's benefits from taking part in activation programs are greater than the costs, and there is a risk that many will choose inactivity, hard paternalism recommends mandatory programs. According to the logic of soft paternalism, mandatory programs should ideally include only those with a failing choice capacity. However, such targeting is extremely difficult, if at all possible, to achieve because it would require reliable testing of each individual recipient so as not to be arbitrary. Otherwise, the inclusion in a mandatory program must only be presumptive and must be cancelled as soon as there is no doubt about a recipient's judgmental capacity or will power. This means that soft paternalism must either make a problematic negative judgment about the choice capacity of welfare recipients overall or turn to a case-by-case-procedure for determining whether individual welfare recipients fall below a choice capacity threshold. If the latter individualizing approach is chosen, soft paternalism may be quite intrusive because one has to find out whether an individual choice is the result of judgmental failures or lack of self-control. That such investigations can be experienced as demeaning to welfare recipients has been a recurrent issue in debates on the design of welfare programs (Titmuss, 1968, Rothstein, 1998). Moreover, practicing soft paternalism will involve an extensive use of discretion by welfare administrators with the associated risk of arbitrary and unequal treatment (Goodin, 1986, Molander, 2016). These problematic aspects of a case-by-case procedure would worry soft paternalists cum liberals, and they would have to think about accountability measures that can counteract them.

Thus, soft paternalism does not seem well-suited to defend welfare conditionality. It does not provide a clear answer to the question of whether it is acceptable to 'throttle' welfare recipients or not. Whether mandatory activation is legitimate or not depends on assumptions about choice capacity that are

difficult to prove. If a soft paternalist would find mandatory activation permissible, he might recommend selective measures trying to target only those who meet certain choice capacity criteria for paternalistic interventions.

One consideration that may be of importance to a soft paternalist's position on welfare conditionality is that it blocks an option (non-participation) for a person who has been guaranteed a right to income support. A negative right is a right to be "selfish, deranged, eccentric, irresponsible, provocative, excessive, irresponsible, monomaniacal etc." (Wellmer, 1993:39) One is not obliged to justify one's actions to others, only to respect their corresponding right to freedom. A soft paternalist could argue that this does not apply to able-bodied persons who live on public benefits. The threshold for interference with self-regarding conduct is lower, and the state, who guarantees income support as a matter of right, may also have a *prima facie* right to impose certain behavioral requirements: for example, to demand participation in work-related programs. Whether to take part in such programs is not a choice that falls within the welfare recipients' legitimate sphere of control. A soft paternalist who reasons this way could recommend mandatory programs.

Hard paternalism is more straightforward. It defends welfare conditionality if there are reasons to assume that mandatory activation, as a general policy, would make welfare recipients better off. It is non-perfectionist if it avoids references to ideas of the good life and argues that welfare recipients are 'thorrered' only to do something they have instrumental reasons to do, regardless of what they otherwise want to achieve. Moreover, it does not involve the complicated and problematic considerations about choice capacity that characterize soft paternalism.

d) Assessing hard paternalism

Although non-perfectionist hard paternalism avoids some problems that accompany soft paternalism, making a convincing paternalistic case of this kind for welfare conditionality is not an easy task. To begin with, there are three conditions that must be fulfilled to have a paternalistic case for mandatory activation of welfare recipients (cf. Conly, 2013:250–52). First, activation must, on average, have a positive effect on the prospect of getting ordinary work. Second, getting a job must provide benefits that surpass the costs of participating in activation programs, evaluated by the recipients' own standards. Third, with voluntary programs, some of those who would benefit would not participate.

Regarding the first condition, there are two reasons why activation requirements may increase future employment among benefit recipients. The positive version is that activation provides skills and knowledge, making participants more employable. The negative version is that activation is considered to be a hassle or a burden that increases the job search activity of those who are on benefits. Note that a paternalist may support mandatory activation even if

the employability effects come via increased searching. However, the extensive empirical literature on the evidence on the employment effects of activation is far from clear, and the long-term labor market outcomes are ambiguous (Molander and Torsvik, 2015).

With regard to the second condition, the problem is this: How can we know that (efficient) mandatory programs will make welfare recipients better off, according to their own standards, and that they will find the costs of participation (the hassle, the reduced leisure time) to be worth it? Regarding the third condition, there are, in principle, several potential reasons why individuals would not participate in a program that benefits them, such as ignorance, akrasia, etc. However, it is very hard to provide solid evidence for this claim. It is not enough to show that, with voluntary programs, not everyone will participate, since, evidently, some of the welfare recipients have no prospects of getting into work.

These three conditions are necessary for having a paternalistic case for mandatory activation. However, they are not sufficient since a justification for this policy must also answer objections based on competing normative concerns. Let us mention three. First, the paternalist must answer the autonomy-based objection against paternalism: thoffer paternalism is wrong because it does not respect its targets as autonomous persons who are capable of reasoning and deciding about their own wellbeing (MacKay, 2019). It is not disrespectful or insulting, per se, to tell people that they are judging and acting imprudently, but to impose a choice on them because they do not understand their own best is something else. The paternalist must argue that preventing people from doing harm to the self is, in principle, not different from preventing people from doing harm to others. In both cases, the question is: What should fall within a person's legitimate sphere of control, and why should it be more disrespectful to exclude harmful self-regarding conduct from this sphere than excluding harmful other-regarding actions? (Hanna, 2018b:73). Whatever one may think of this matter, it is not obvious that the choice to decline participation in an activation program has such serious negative consequences for welfare recipients that, for paternalistic reasons, it does not belong to the sphere of legitimate personal control. Second, a thoffer policy does not target only individuals who are unwilling to participate. It will also target persons who cannot work, and persons who are willing to participate. Forcing the first group is unfair, while forcing the latter group is unnecessary and a waste of resources. Third, activation programs may be expensive to the welfare state and burden other citizens as taxpayers. A pure paternalist may argue that mandatory programs should be chosen politically anyway, because they help disadvantaged individuals to better their opportunities for welfare. However, program costs, cost-efficiency, and prioritization of public expenditures are questions that can hardly be avoided by any public justification of mandatory programs. Moreover, to justify to other citizens that they

should contribute to the funding of these programs, paternalistic reasons are not enough. The paternalist must also appeal to an obligation of beneficence toward welfare recipients.

Coda

By connecting mandatory activation with the philosophy of paternalism, we have tried to provide a better understanding of what is involved in defending welfare conditionality for paternalistic reasons. When we conclude that hard paternalism without a perfectionistic grounding is better suited than soft paternalism to justify mandatory activation, this does not mean that we think hard paternalism provides a convincing case for welfare conditionality. We only claim that if you want to provide a paternalistic defense of mandatory activation, then hard paternalism is better suited. However, the problems it brings are of such magnitude that it falls short as a convincing justification of mandatory activation. Other justifications may turn out to be more successful, but that remains to be seen.¹¹

Notes

- 1 We want to thank Axel West Pedersen, Johannes Lindvall, Lars Inge Terum and Cathrine Holst and two anonymous reviewers for their constructive comments and suggestions.
- 2 According to OECD activation policies aim “to bring more people into the effective labor force, to counteract the potentially negative effects of unemployment and related benefits on work incentives by enforcing their conditionality on active job search and participation in measures to improve employability, and to manage employment services and other labor market measures so that they effectively promote and assist the return to work” (OECD, 2013:132).
- 3 For overviews see Coons and Weber, 2013; Grill and Hanna, 2018; Le Grand and New, 2015: 7–38.
- 4 See, for example, Dworkin, 1972 and 2013; Feinberg, 1986; Le Grand & New, 2015; Scoccia, 2018; Hanna, 2018a; Wall, 2018.
- 5 Here we have linked perfectionism to paternalism, but note that anti-paternalism can also have a perfectionistic justification. See John Stuart Mill on individuality as “one of the leading essentials of well-being” (Mill, 1989: 57). Liberal neutralists would consider this kind of anti-paternalism as too “comprehensive”.
- 6 Cf. John Rawls who assumes that the parties in the original position will want to ‘protect themselves against their own irrational inclinations by consenting to a scheme of penalties that give the sufficient motive to avoid foolish actions and by accepting certain impositions designed to undo the unfortunate consequences of their imprudent behavior’ (Rawls, 1971:249).
- 7 Cf Isaiah Berlin: “to be free to choose, and not be chosen for, is an inalienable ingredient in what makes human beings human” (Berlin, 1969:lx).
- 8 Cf. the dialogue between “crazy” and “lazy” in van Parijs, 1991.
- 9 See for example Kojève in his *Introduction to the reading of Hegel* (Kojève, 1980 [1947]:25): “The product of work is the worker’s production. It is the realization of his project, of his idea; hence, it is he that is realized in and by this product, and consequently he

contemplates himself when he contemplates it . . . Therefore, it is by work, and only by work, that man realizes himself objectively as man.’ For a more modest but still demanding version of the idea of self-realization, see Elster, 1986.

- 10 By public justification, we mean acts of reason giving that are public in three respects: the reasons pertain to “public things,” such as laws; they are “presented in the public or by citizens acting publicly”; and they are addressed to “a public at large characterized by pluralism” (Chambers, 2010, 894). The last aspect is the crucial one. A public justification must be viable under conditions of disagreements between citizens embracing often incommensurable “comprehensive doctrines,” meaning conceptions of the meaning, value, and purpose of human life (see Rawls, 1993: 58–65).
- 11 See Molander and Torsvik, 2015.

References

- Berlin, I. (1969), *Four Essays on Liberty*. Oxford: Oxford University Press.
- Chambers, S. (2010), ‘Theories of public justification’, *Philosophy Compass*, 5(11), 893–903.
- Conly, S. (2013), *Against Autonomy, Justifying Coercive Paternalism*. Cambridge: Cambridge University Press.
- Coons, C. and Weber, M. (eds.) (2013), *Paternalism. Theory and Practice*. Cambridge: Cambridge University Press.
- Darwell, S. (2006), ‘The Value of Autonomy and the Autonomy of Will’. *Ethics* 116: 2, 261–284
- Dunn, A. (2014), *Rethinking Unemployment. Beyond the ‘Quasi-Titmuss’ Paradigm*. Houndmills: Palgrave Macmillan.
- Dworkin, G. (1972), ‘Paternalism’. *The Monist*, Vol. 56, Issue 1, 64–84.
- Dworkin, G. (1991), *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press.
- Dworkin, G. (2013), ‘Defining Paternalism’. In: Coons, C. and M. Weber (eds.), *Paternalism. Theory and Practice*. Cambridge: Cambridge University Press.
- Dwyer, J. (2020), ‘Questions of conduct and social justice: the ethics of welfare conditionality within UK social security’. In: Eleveld, A., Kampen, T., Arts, J. (eds.), *Welfare to Work in Contemporary European Welfare States*. Bristol: Policy Press.
- Eichhorst, W., Kaufman, O. and Konle-Seidl, R. (eds.) (2010), *Bringing the jobless into work? Experiences with activation schemes in Europe and the US*. Berlin: Springer.
- Elster, J. (1986), ‘Self-realisation in work and politics: the Marxist conception of the good life.’ *Social Philosophy and Policy* 3, 97–126.
- Eriksen, A. and Molander, A. (2019). Welfare Reform and Public Justification. *Policy Studies*, Vol. 40, Issue 6.
- Feinberg, J. (1986), *Harm to Self*. Oxford: Oxford University Press.
- Goodin, R.E. (1986). Welfare, Rights and Discretion. *Oxford Journal of Legal Studies*, Vol. 6, 2.
- Goodin, R.E. (1998), ‘Social welfare as a collective social responsibility’. In: Schmitz, D. and Goodin, R.E. (eds.), *Social Welfare and Individual Responsibility*. Cambridge: Cambridge University Press.
- Grill, K. and Hanna, J. (eds.) (2018), *The Routledge Handbook on the Philosophy of Paternalism*. London & New York: Routledge.
- Hanna, J. (2018a), ‘Hard and Soft Paternalism’. In Grill, K. and J. Hanna (eds.), *The Routledge Handbook on the Philosophy of Paternalism*. London & New York: Routledge.
- Hanna, J. (2018b), *In Our Best Interest. A Defence of Paternalism*. Oxford: Oxford University Press.
- Henning, C. (2009), ‘Liberalism, Perfectionism and Workfare’. *Analyse&Kritik*. Jahrgang 31, Heft 1, 159–180
- Höfle, V. (1997), *Moral und Politik: Grundlagen einer Politischen Ethik für das 21. Jahrhundert*, München: Beck.

- Jahoda, M. (1982), *Employment and unemployment: A social-psychological analysis*. Cambridge, England: Cambridge University Press.
- Kojève, A. (1980 [1947]), *Introduction to the Reading of Hegel. Lectures on the Phenomenology of the Spirit assembled by Raymond Queneau*. Edited by Harold Bloom. Ithaca and London: Cornell University Press.
- Le Grand, J. and New, B. (2015), *Government Paternalism. Nanny State or Helpful Friend?* Princeton: Princeton University.
- MacKay, D. (2019), 'Basic Income, Cash Transfers, and Welfare State Paternalism'. *The Journal of Political Philosophy*, Vol. 27, Number 4, 2019, 422–447
- Marx, K. (1844/1975), 'Economic and Philosophical Manuscripts', in Marx and Engels, *Collected Works*, Vol. 3. London: Lawrence & Wishart.
- Mead, L. (1997), 'The Rise of Paternalism'. In *The New Paternalism. Supervisory Approaches to Poverty*. Ed. by Lawrence Mead. Washington D.C.: Brookings Institution Press.
- Mead, L. (2005), 'Welfare Reform and Citizenship'. In Mead, L., Beem, C. (eds.), *Welfare Reform and Political Theory*, New York: Russell Sage Foundation.
- Mestan, K. (2016). 'Paternalism in Australian Welfare Policy'. *Australian Journal of Social Issues*, Vol. 49 (1).
- Mill, J. S. (1989 [1859]). In J.S. Mill, *On Liberty and other writings*. Ed. By Stefan Collini. Cambridge: Cambridge University Press.
- Molander, A. and Torsvik, G. (2015), 'Getting People into Work: What (if Anything) Can Justify Mandatory Activation of Welfare Recipients?' *Journal of Applied Philosophy*, Vol 32, Issue 4.
- Molander, A. (2016). *Discretion in the Welfare State: Social Rights and Professional Judgement*. London: Routledge.
- OECD (2013). *Employment outlook 2013*. Paris: OECD Publishing.
- Quong, J. (2011), *Liberalism without Perfection*. Oxford: Oxford University Press.
- Raz, J. (1986), *The Morality of Freedom*. Oxford: Clarendon Press.
- Rawls, J. (1971), *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls, J. (1993), *Political Liberalism*. New York: Columbia University Press.
- Rothstein, B. (1998), *Just Institutions Matter. The Moral and Political Logic of the Universal Welfare State*. Cambridge: Cambridge University Press.
- Scoccia, D. (2018), 'The Concept of Paternalism'. In Grill, K. and J. Hanna (eds.), *The Routledge Handbook on the Philosophy of Paternalism*. London & New York: Routledge.
- Shapiro, D. (2007), *Is the Welfare State Justified?* Cambridge: Cambridge University Press.
- Shiffrin, S. (2005), 'Paternalism, unconscionability doctrine, and accommodation'. *Philosophy & Public Affairs* 29, 3: 205–250.
- Soss, J., Fording, R.C. and Schram, S.F. (2011), *Disciplining the Poor: Neoliberal Paternalism and the Persistent Power of Race*. Chicago: University of Chicago Press.
- Steiner, H. (1974–75). Individual liberty. *Proceedings of the Aristotelian Society*, 75: 33–50
- Sugden, R. (2019), *The Community of Advantage. A Behavioural Economist's Defence of the Market*. Oxford: Oxford University Press.
- Sunstein, C. R. (2014), *Why Nudge? The Politics of Libertarian Paternalism*. New Haven & London: Yale University Press.
- Sunstein, C. R. (2016), *The Ethics of Influence. Government in the Age of Behavioral Science*. Cambridge: Cambridge University Press.
- Swift, A. and White, S. (2008), 'Political theory, Social science and Real Politics'. In: Leopold, D. and Stears, M. (eds.), *Political Theory: Methods and Approaches*. Oxford: Oxford University Press, 49–69.
- Taylor, M. (1982), *Community, Anarchy and Liberty*. Cambridge: Cambridge University Press.
- Thaler, R.H. and Sunstein, C. R. (2008), *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- Titmuss, R. (1968), *Commitment to Welfare*. London: George Allen Unwin.
- Van Parijs, P. (1991), 'Why Surfers Should be Fed: The Liberal Case for an Unconditional Basic Income'. *Philosophy & Public Affairs*, Vol. 20, No. 2, 101–131

- Van Parijs, P. and Vanderborght, Y.Y. (2017), *Basic Income. A Radical Proposal for a Free Society and a Sane Economy*, Cambridge (MA) and London: Harvard University Press.
- Wall, S. (1998), *Liberalism, Perfectionism and Restraint*, Cambridge: Cambridge University Press.
- Wall, S. (2018), 'Perfectionism and paternalism'. In: *The Routledge Handbook of the Philosophy of Paternalism*. Ed. by K. Grill and J. Hanna. New York: Routledge.
- Watts, B. and Fitzpatrick, S. (2018), *Welfare Conditionality*. London: Routledge
- Wellmer, A. (1993), *Endspiele: Die unversöhnliche Moderne. Essays und Vorträge*. Frankfurt am Main: Suhrkamp. English translation: *Endgames: The Irreconcilable Nature of Modernity*. Cambridge Mass.: MIT Press 2000.
- Wertheimer, A. (1987), *Coercion*. Princeton: Princeton University Press.
- White, M. D. (2011), *Kantian Ethics and Economics. Autonomy, Dignity and Character*. Stanford: Stanford University Press.
- White, S. (2005), 'Is Conditionality Illiberal?'. In: *Welfare Reform and Political Theory*. Ed. by Lawrence H. Mead and Christopher Beam. New York: Russel Sage Foundation.
- White, S. (2008), *The Civic Minimum. On the Rights and Obligations of Economic Citizenship*. Oxford: Oxford University Press.
- Wolff, J. (2004), 'Training, Perfectionism and Fairness'. *Journal of Applied Philosophy*, 21(3), 285–95.