



Dementia, Authenticity, and the Moral Weight of Preferences

Sigurd Lauridsen · Somogy Varga

Received: 24 November 2025 / Accepted: 25 March 2026 / Published online: 2 April 2026
© The Author(s) 2026

Abstract The moral weight of preferences expressed by people with advanced dementia remains a persistent challenge for clinical ethics, in part because healthcare professionals (HCPs) often treat some such preferences as still deserving respect even while acknowledging that decision-making capacity is compromised. Drawing on dilemmas discussed in a 2023 ethics workshop in a Danish nursing home, we argue that HCPs differential responses are plausibly tracking a distinction between disorder-driven preferences and *authentic* preferences that remain genuinely expressive of the person. We defend a *thin*, procedure-externalist conception of authenticity on which preferences can be authentic even without higher-order critical reflection, and we argue that such authenticity can ground defeasible normative reasons for compliance even without autonomy. We address three challenges and note converging experimental-bioethics findings suggesting that stakeholders' judgments in dementia cases are sensitive to authenticity considerations. Given persistent uncertainty, we propose three defeasible, clinically usable indicators for presuming

authenticity: (i) stability and minimal prioritisation in preference patterns, (ii) domain-specific minimal ("spherical") comprehension, and (iii) absence of salient distortion or undue influence. We apply these indicators to the workshop cases and illustrate how the framework can guide practice while leaving room for justified override in high-stakes or distorted cases. Overall, the framework extends current debates in neuroethics both rationalises a common source of clinical hesitation about capacity-based decision rules and offers a principled, action-guiding approach to care decisions.

Keywords Authenticity · Autonomy · Dementia · Decision-making capacity · Moral authority · Neuroethics

The question of the moral weight of preferences expressed by people with advanced dementia continues to pose a challenge in both clinical ethics and practice. This problem was vividly illustrated during a 2023 ethics workshop with healthcare professionals (HCPs) at a Danish nursing home. During the workshop, a heated debate emerged over residents' rights to self-determination in dementia care [1, 2]. The discussion was sparked by a literary case from Danish author Kirsten Thorup's book *No Man's Land*¹

S. Lauridsen (✉)
The National Institute of Public Health, University
of Southern Denmark, Odense, Denmark
e-mail: sila@sdu.dk

S. Varga
Department of Philosophy and the History of Ideas,
Aarhus University, Aarhus, Denmark
e-mail: varga@cas.au.dk

¹ Our own translation of the title from Danish to English.

(2004), which, among other things, explores whether a person with dementia has the right to die if they so choose. The case featured an elderly man in a nursing home refusing food and water, leading to an intense exchange among the 25 HCPs present, all of whom worked with people living with dementia. As the debate unfolded, participants revealed conflicting views on decision-making capacity in advanced dementia. While they agreed that individuals with severe dementia lacked full competency in care decisions, they also expressed a strong commitment to respecting their autonomy whenever possible. This tension surfaced in three dilemmas drawn from their clinical experience.

The first dilemma involved a lifelong vegetarian with dementia who suddenly craved meat. Most participants hesitated to serve it, especially if the residents had documented their preference against meat. However, if the residents persisted, many would eventually accommodate the new desire. The second case concerned a Muslim resident who requested pork. Here, participants (primarily Muslim HCPs) were more reluctant, feeling it was inappropriate to serve pork regardless of the resident's desire. Unlike the vegetarian case, no clear consensus emerged. The final dilemma echoed Thorup's case: residents refusing food and water. Many HCPs viewed force-feeding as undignified and, unlike in the previous cases, there was unanimous agreement that such refusals should be honoured. One participant described a resident experiencing "a beautiful moment of clarity" when choosing to stop eating and drinking. When asked why they readily respected a resident's refusal to eat and drink but hesitated over dietary choices, participants struggled to articulate a principled rationale: they remained confident in the distinction despite lacking an explicit justificatory framework.

Interestingly, the participants' responses suggest that their practical reasoning was not guided solely by assessments of decision-making capacity, nor by straightforward appeals to best interests or to previously expressed preferences. Instead, they appeared to rely on a distinction between preferences that could be dismissed as symptoms of disease and preferences that were experienced as *authentic* (i.e., genuinely expressive of the person), even in advanced dementia, and thus as having moral significance that calls for respect. This is noteworthy not merely because it suggests that an important normative distinction is doing

practical work without being made explicit, but also because it appears to run counter to a standard line of reasoning, which rests on two key assumptions: (i) that autonomy-related capacity with respect to a given decision is both a necessary and sufficient condition for full decisional authority; and (ii) that the absence of such capacity entails the absence of any entitlement to decisional authority. In particular, the HCPs judgments seem to conflict with assumption (ii): even while acknowledging the absence of decision-making capacity, they continue to treat some expressed preferences as carrying moral weight.

But this raises a concern. While the pattern of judgments described above suggests that HCPs are tracking a morally relevant, authenticity-related difference, it remains unclear whether the distinction, as deployed in practice, is normatively defensible. In particular, can HCPs have authenticity-based reasons to respect certain expressed preferences even when decision-making capacity is absent? Treating these HCPs' judgments as an illustrative starting point, the central claim of this paper is that although people with advanced dementia typically lack autonomy and decision-making capacity, they may nevertheless retain a capacity for authenticity, understood as the ability to express preferences that are, in an important sense, genuinely their own. This grounds a limited but morally significant form of decisional authority, such that HCPs may, in certain circumstances, have *authenticity-based reasons* to comply with individuals' preferences. By developing this claim, we aim both to explain the distinction that HCPs *could* be tracking in practice and to contribute to more principled reflection on similar dilemmas in clinical care.

The paper proceeds as follows. Section "[Decision-making capacity](#)" outlines the standard conception of decision-making capacity and its role as a clinical proxy for autonomy, with particular attention to why advanced dementia is typically taken to undermine capacity while still leaving clinicians hesitant in practice. Section "[Autonomy and Authenticity](#)" develops the conceptual setup: it introduces autonomy and authenticity via Christman's account (§2), argues for a thin, procedure-externalist conception of authenticity that can persist without higher-order critical reflection (2.1), and briefly connects this distinction to findings from experimental philosophical bioethics ("bioxphi") suggesting that stakeholders' judgments in dementia cases are sensitive to authenticity

("true self") considerations over and above capacity (2.2). Section "[Objections and replies](#)" addresses three objections and uses them to motivate the need for defeasible indicators rather than a decisive test of authenticity. Section "[Situating the view in the literature: two routes](#)" situates our proposal in the dementia ethics literature, contrasting it with valuing-based and thick-authenticity routes and clarifying what our approach adds. Section "[Conditions for adopting authentic preferences](#)" then develops three defeasible, clinically usable indicators (consistency, minimal comprehension, and absence of salient distortion or undue influence) that can justify presuming authenticity under conditions of uncertainty. Section "[Implications and real-live cases](#)" applies the framework to the workshop cases and shows how it can help make sense of the observed pattern of clinical judgments while acknowledging potential confounds. Section "[Further implications](#)" draws broader normative implications, including how authenticity-based reasons relate to welfare- and autonomy-based considerations, before concluding with limitations and directions for future empirical and philosophical work.

Decision-making Capacity

Across severe mental disorders and neurocognitive disorders such as dementia, a recurring concern is whether an individual's expressed preferences genuinely reflect autonomous agency or are instead products of the disorder itself. When preferences appear imprudent, self-harming, or sharply inconsistent with a person's prior values, this concern translates into uncertainty about whether they ought to be respected.

Under ordinary circumstances, persons are widely assumed to have a default right to self-determination, which we are required to respect. John Rawls, for instance, justifies this right by arguing that persons are "self-authenticating sources of valid claims"[3], underwriting the moral requirement to take their expressed choices as authoritative. By contrast, it is commonly held that individuals who lack the capacity for autonomy do not possess the same determinate right to self-determination[4]. In clinical contexts, this asymmetry grounds the permissibility of weak paternalism: acting in a person's best interests even when doing so conflicts with what the person presently wants, in order to prevent refusals of treatment

that would have been accepted under adequate understanding, or consent to interventions that would have been declined [5].

Historically, there has been considerable debate about how certain medical conditions affect individuals' decision-making capacity and thereby when paternalism is justified. This debate has given rise to a range of positions, including two opposing extremes. On the one hand, views that treat certain conditions *en bloc* as impairing decision-making capacity to such an extent that affected individuals should be deemed legally incompetent to make treatment choices. On the other hand, views that maintain that these patients generally retain adequate decision-making abilities and should therefore retain their legal right to make binding treatment decisions[6]. Steering between these somewhat extreme positions, various measurement instruments have been developed to assess decision-making capacity, among them the MacCAT-T, widely regarded as the gold standard for assessing decision-making capacity, including in dementia care[7].

Treating decision-making capacity as a proxy for autonomy[8], the scale specifies four interrelated abilities that together establish such capacity. First, an individual must be able to express a choice, meaning they can articulate a clear and consistent treatment decision, which need not conform to external standards of rationality but must display stability over time. Second, they must demonstrate understanding of the information relevant to the decision, including the nature of their condition, the proposed interventions, the associated risks and benefits, available alternatives, and the possibility of foregoing treatment altogether. Third, they must be able to appreciate how this information bears on their own situation, recognising the likely consequences of accepting or refusing treatment and acknowledging its implications for their health and life. Finally, they must demonstrate the ability to reason about the available options by comparing risks and benefits and provide a rationale for their decision, even if the outcome deviates from what others might regard as the optimal choice.

On its face, assessments of decision-making capacity in individuals with advanced dementia seem relatively straightforward: they fail to meet the criteria specified in instruments such as the MacCAT-T, particularly with respect to understanding alternatives, weighing reasons, and providing a rationale for

decisions[9]. Nonetheless, despite the authoritative status of these instruments, HCPs often experience uncertainty and hesitation, finding that patients resist clear classification as simply having or lacking capacity in relation to specific preferences and care decisions. Caregivers sometimes, for instance, refuse to force treatment on people living with dementia, even when they believe treatment to be in the individual's best interest and psychiatric assessments indicate a lack of decision-making capacity[10]. Thus, a grey area arises in which HCPs are unsure whether a person is making an appropriate care choice, as evidenced by the workshop cases we described, regardless of how they score on the relevant scales.

The persistent hesitation to rely on capacity assessments alone raises a deeper question: which aspects of autonomy are failing in advanced dementia, and does anything remain that can still be authoritative in grounding at least *prima facie* reasons for respecting patients' preferences? To address this question, we begin by drawing on a well-known account of autonomy. This framework allows us to clarify which components of autonomy are typically missing in advanced dementia, while also helping to disentangle other dimensions of agency that may persist despite cognitive impairment. On this basis, we argue that although people with advanced dementia lack autonomy in the full sense, they may nevertheless retain authentic preferences that are at least *prima facie* worthy of respect.

Autonomy and Authenticity

It seems to be a uniting feature of almost all, if not all, plausible theories of autonomy that a necessary condition for a person to have the capacity for autonomy is the ability to make expressions that genuinely display her values. In the words of Christman, autonomy is "the ability of the person to guide her life from her own perspective, rather than being manipulated by others or being forced into a particular path by surreptitious or irresistible forces"[11] (p. 278). Christman proposes that preferences must satisfy two broad conditions in order to count as autonomous. The *competence condition* requires that individuals possess the cognitive capacities necessary to form effective intentions on the basis of the preferences in question, as well as a general capacity to reflect critically

on those preferences. This requirement aligns with a near consensus in the literature that, for a preference or other motivational state to be genuinely expressive of an individual, the individual must have sufficiently intact psychological mechanisms to allow for reflection upon, and potential revision of, their beliefs and desires [12]. The *authenticity condition* requires that preferences be adopted through a process from which the person would not have felt alienated had she been able to reflect freely on the preference's content and origins. This requirement is plausible because it strikes a balance: it neither sets the bar so high that ordinary, unreflectively acquired values are excluded, nor so low that preferences developed via disease or manipulation automatically count as autonomous. This aligns with the intuition that when a person's preferences are the product of a mental disorder, they do not genuinely express the person as such, but are instead imposed upon them by an external force[12].²

While the dementia literature largely agrees that people with advanced dementia continue to have preferences, in the minimal sense that there are things the person likes better or would choose over alternatives[13], it also agrees that they lack the ability to reflect critically on those preferences. This means that they do not meet the competence condition: they lack the cognitive capacities necessary to reflect critically on their preferences and to form effective intentions on the basis of those preferences. Along similar lines, Karin Jongsma and colleagues argue that because individuals with severe dementia are unable to adopt preferences self-reflexively, the preferences they develop in this state should not be straightforwardly trusted as expressions of autonomous agency[14]. Considerations of this kind also underlie the widespread endorsement of advance directives by philosophers, healthcare systems, and patient organisations[15]. Advance directives allow individuals to make informed decisions about significant aspects of

² On this view, autonomy is context-relative rather than a global status. For example, an individual may be autonomous with respect to their political preferences but not with respect to their religious commitments, depending on the historical processes through which those preferences arose and on whether the person would endorse them upon reflection. This context-sensitivity makes this account apt for discussions of dementia, where impairment is often uneven or domain-specific.

their future care at a time when they still possess the capacity for autonomous decision-making, thereby extending the authority of earlier autonomous preferences into periods in which such autonomy is no longer available.

Some may be inclined to conclude that people with advanced dementia also fail to meet the authenticity condition, on the grounds that they lack the reflective capacities required to determine whether they are alienated from their preferences. However, as we shall argue in the next sections, this conclusion does not necessarily follow. What advanced dementia rules out is autonomy in the full sense, not necessarily the authenticity of every expressed preference, and such authenticity may, in turn, ground a limited but morally significant form of decisional authority.³

Procedure Internalism, Externalism, and "thin" Authenticity

Can preferences be authentic without critical reflection? On a procedure-internal reading, they cannot: authenticity is a product of critical reflection, not a property preferences can possess independently of it. This view is difficult to reconcile with the *Real-Life Dilemma* of autonomy theories, since it implies that many ordinary preferences are neither autonomous nor non-autonomous simply because they have never been reflected upon.⁴ As Michael Quante notes, demanding reflection on every preference would make autonomy "too strong an ideal for human beings," since "not exercising the capacities necessary for personal autonomy cannot be sufficient for not being autonomous at all"[16] (p. 598). Christman

himself stresses that "reflection [e.g. on preferences] should be hypothetical, in that few of us have reflected on all those aspects of ourselves that, nevertheless are thoroughly authentic and free formed"[11] (p. 279).

The alternative view, procedure-externalism, holds that preferences can be authentic from the outset, independently of any actual critical reflection. Thus, a preference is authentic if it would withstand the relevant counterfactual test: were the agent to engage in sustained critical reflection on the preference under appropriate conditions (i.e., in light of its formation history and without reflection-distorting influences) she would not be alienated from it. Authenticity, on this account, is therefore not produced by reflection. Instead, reflection functions as a hypothetical standard for assessing whether the preference is genuinely the agent's own. Of course, we can distinguish between preferences whose status is epistemically unsettled (because the agent has not reflected on them) and preferences whose status is settled (because the agent has engaged in critical reflection and found herself non-alienated). But on this view, reflection does not create authenticity, but merely provides evidence of how the relevant counterfactual test would come out.

Procedure-externalism thus supports a *thin notion of authenticity*: preferences can count as authentic without presupposing the richer, self-evaluative or identity-constituting capacities characteristic of "thick" authenticity[17, 18]. On this view, a preference may be authentic even when the higher-order reflective capacities required for autonomy are absent. Thus, preferences can be authentic without critical reflection: although their status may be epistemically unsettled, some preferences will nonetheless satisfy the authenticity condition. This implies that even when people with advanced dementia lack autonomy in the full, competence-involving sense, it can still be possible for them to express authentic preferences, even if she is cognitively prohibited from reflectively identifying them as such.

Having argued that preferences can be authentic without critical reflection, a natural follow-up question is whether such preferences can carry *prima facie* moral weight and so ground authenticity-based reasons for respect and compliance, even when decision-making capacity is absent. To explain, we can start from the well-established distinction between *care* for persons and *respect* for persons[19]. On this view,

³ Some may prefer to describe our target property as a "thin" form of autonomy. We avoid that label because "autonomy" is commonly associated with more demanding capacities (robust understanding, deliberation, and reflective endorsement) and is typically taken to ground full decisional authority. Our aim is narrower: to identify conditions under which preferences expressed below standard autonomy and capacity thresholds can nonetheless have defeasible normative weight, without suggesting that they carry the same authority as autonomous choices.

⁴ We should note that not everyone sees this implication as a cost. For example, both Raz and Frankfurt treats autonomy as an *ideal* rather than a status that attaches to every ordinary preference, so it is unsurprising on his view if many unreflected preferences fall outside autonomy. For a similar position in the context of authenticity, see Varga (2012).

to care for someone is to have largely agent-neutral reasons to promote what is good for them, whereas to respect someone is to have reasons to relate to them in a way that takes seriously their *evaluative standpoint*, roughly, what they themselves value from their point of view. Respect, in this sense, involves treating people as subjects who occupy a perspective from which things matter. One way of doing so is to treat their expressed preferences as *prima facie* reason-giving: other things equal, they count in favour of compliance. But a preference can play this role only if it is authentic, i.e., if it is genuinely grounded in the person's evaluative standpoint rather than explainable as a product of distortion (e.g., disorder-driven influences, coercion, manipulation). Autonomy strengthens decisional authority by adding demanding competence and reflective control, but those capacities, at least on the thin account, are not required for a preference to be authentic. Thus, even in people who do not meet thresholds for autonomy and decision-making-capacity, authentic preferences can retain defeasible normative significance.

Experimental Findings

The preceding section motivates, at the conceptual level, distinguishing two dimensions that are often run together in clinical discussions: competence (roughly, the cognitive and reflective abilities associated with autonomy) and authenticity (roughly, whether a preference is genuinely the agent's own). This leaves room for the possibility that some preferences satisfy a thin authenticity condition and retain defeasible normative weight even when autonomy capacity is absent. Importantly, this separation is not only theoretically available. It is also reflected in how stakeholders commonly reason about dementia cases, where judgments about what ought to be done track authenticity considerations over and above capacity. In the following, we briefly highlight this experimental evidence, which lends some initial support to our claim that authenticity can retain defeasible normative significance.

Experimental philosophical bioethics ("bioxphi") is a branch of experimental philosophy that brings controlled vignette studies and related methods from psychology to bear on questions in bioethics[20]. In particular, a growing subset of bioxphi focuses on dementia and advance-care planning, using experimentally controlled vignettes to study how people

weigh capacity, advance directives, and authenticity-like considerations when deciding what ought to be respected. Rather than treating stakeholders' intuitions as background assumptions, bioxphi sets out to examine which factors (e.g., autonomy, capacity, authenticity) systematically influence judgments about which preferences ought to be respected.

Across this literature, a consistent result is that laypeople's judgments about whether a patient's (medical) preference should be respected contrasts with the traditional (Dworkinian) approaches to advance directives in which decisional authority tracks capacity: once a patient, at t_2 , falls below the relevant threshold of autonomy, her current preferences are typically taken to lose their claim to be followed, and earlier capacitated choices from t_1 (from advance directives or surrogate judgment) take priority. In contrast, laypeople's judgments are shaped less by autonomy than by whether the preference is seen as authentic in the sense of expressing who the patient "really is"[21]. Even when laypeople judge that a patient with dementia lacks autonomy or decision-making capacity, they may still treat some of the patient's expressed preferences as authentically her own (i.e., as reflecting her "true self") and therefore as carrying moral weight. For example, in a vignette study involving a dementia patient described as lacking decision-making capacity, participants who nonetheless judged her to remain her "true self" were more inclined to honor her current preference, even when doing so conflicted with her advance directive[22]. Importantly, this work should not be read as endorsing any general rule that a person's current preferences at t_2 automatically override an advance directive made at t_1 . Rather, it suggests that preferences formed under conditions of autonomy (e.g., in an advance directive) and authenticity considerations regarding current expressed preferences after autonomy is lost can function as competing sources of normative weight.

Overall, while these studies are not decisive for the normative question, they show that the autonomy–authenticity distinction is not merely a theoretical possibility: stakeholders' judgments in dementia cases often treat authenticity-like considerations as normatively salient in practice. This, in turn, motivates developing a defeasible, clinically usable framework for assessing when an expressed preference is plausibly authentic, and thus merits respect. We develop this framework in Section "[Conditions](#)

for adopting authentic preferences", after addressing objections and further situating our account.

Objections and Replies

Before proceeding, we briefly flag two reservations about whether a counterfactual non-alienation test can do determinate normative work in advanced dementia. First, on the *Counterfactual Applicability Objection*, once we ask how to understand the idealizing antecedent (roughly, if the patient were to engage in undistorted critical reflection) the test threatens triviality or subject-switching. Second, on the *Preference-Formation Objection*, even on a thin view, one might doubt that advanced dementia leaves sufficient agential capacity for preferences to be genuinely formed and owned. Third, on the *Past-Self Privileging Objection*, even granting that some present preferences may be authentic, epistemic uncertainty may appear to justify treating pre-dementia preferences as the default guide to what should be respected.

The first objection could start from the claim that, if it is impossible for the patient to critically reflect on one of her preferences, then, at least on a Lewis–Stalnaker-style account, the counterfactual in question would be vacuously true: such conditionals are evaluated by looking to the closest possible worlds, and if there are no such worlds because the antecedent is impossible, the truth condition is satisfied trivially. Thus, if it is impossible for the patient to critically reflect on one of her preferences, which is plausible especially in advanced stages, the counterfactual test would be automatically met, yielding the trivial yet absurd conclusion that *all* of her preferences are authentic.⁵ A first reply is that vacuism about counterpossibles is contested: many reject automatic truth for impossible-antecedent counterfactuals because it makes them indiscriminate.⁶ Moreover, discussions

of counterpossibles typically target antecedents that are metaphysically impossible, not antecedents that are merely nomologically impossible, which need not generate the same vacuity worries. But on that framing, the objection does not undermine our proposal, because we might insist that critical reflection is not impossible in the relevant sense, since there are possible worlds in which the patient simply lacks dementia.

But this response invites a further worry: a dementia-free counterpart may differ so substantially (for instance in her psychology and overall evaluative outlook) that she no longer plausibly counts as the *same* person whose preferences are under consideration. But does our account presuppose a degree of diachronic psychological continuity depending on psychological links (e.g., memory and a stable evaluative outlook) that advanced dementia can undermine? If those links are too weak, the counterfactual risks answering the wrong question, because the only plausible reflective counterpart is no longer the same person whose present preference we are assessing. Alternatively, it risks failing to yield a determinate answer, because there is no principled way to specify which counterfactual continuation counts as the relevant version of *this* person under reflection.

However, the objection is significantly weakened once we clarify that our appeal to counterfactual endorsement presupposes only a minimal diachronic identity akin to narrative or relational accounts[23], that is enough for preference-ascription but does not require robust psychological continuity. Note that this thin diachronic identity is not an extra metaphysical claim but is already built into the kind of practical question at issue here. In other words, it is already presupposed whenever we ask: "Should we respect this person's current preference rather than his earlier preference?" you are already assuming some diachronic relation robust enough for it to be the same bearer of preferences across times (or, at minimum, for their conflict to be intelligible as a conflict within one person's practical standpoint). On such a conception of continuity, counterfactual endorsement is meaningful even in where cognitive capacities are compromised.

A second objection (the *Preference-Formation Objection*) is that even on a thin account, a present preference can count as authentic only if it is the sort of attitude the agent can, in some minimal sense,

⁵ We are grateful to an anonymous reviewer for suggesting this complication.

⁶ For a brief illustration of why some philosophers reject vacuism about counterpossibles, consider: "If Hobbes had squared the circle, mathematicians would be surprised" versus "If Hobbes had squared the circle, he would have ended world hunger" (Kocurek, 2021). While these are not truth-conditionally equivalent (and thus vary in truth value), vacuism predicts that both are true, erasing the difference between conditionals that track relevant dependence and those that are plainly irrelevant.

own. Yet such ownership plausibly requires cognitive capacities to form, sustain, and adopt preferences as one's own, capacities that advanced dementia may compromise[24]. However, the claim that non-autonomous individuals therefore cannot form authentic preferences at all is inadequately supported, since the capacities required for ownership on a thin view are comparatively modest. Many preferences originate under conditions short of autonomy (e.g., in childhood) yet are not thereby inauthentic. That said, lacking or limited autonomy plausibly increases vulnerability to distortion and undue influence, which is why it is better treated as a *risk factor* for inauthentic preference formation. Where minimal capacities are absent, authenticity attributions will be correspondingly weaker or underdetermined, precisely motivating the need for defeasible indicators.

A third objection, (*Past-Self Privileging Objection*) starts by granting that some present preferences in advanced dementia can be authentic in a thin sense, but stresses that we are often not in a position to tell which ones are. Given this epistemic uncertainty, the most defensible practical default, the objection continues, is to privilege the person's pre-dementia preferences (especially as expressed in an advance directive or reconstructed by a surrogate) as the most reliable evidence of what expresses the person's evaluative standpoint and therefore ought to be respected. This line of thought is common in the literature. For example, Stirland et al.[25], who define authenticity broadly as "being true to oneself," note that "If authenticity means being true to oneself and the self is becoming lost, it may seem that people with Alzheimer's dementia can only be authentic to their past self." On this approach, a surrogate (typically someone with intimate knowledge of the person's pre-dementia evaluative outlook) is tasked with reconstructing what the person endorsed prior to decline (*t1*) and assessing whether current expressions (*t2*) cohere with that earlier standpoint[26]; marked divergence counts against treating the present preference as authentic (or as carrying decisional weight), whereas coherence counts in its favour.

This objection can be resisted. Even if advanced dementia increases the risk of distorted preference formation, it does not follow that the person's pre-dementia standpoint (*t1*) is automatically authoritative for what she now wants (*t2*). Under ordinary circumstances, divergence from earlier commitments

is not sufficient to disqualify present preferences as "not genuine": people can and do undergo sometimes radical changes. Moreover, there is specific reason to doubt that *t1* reliably predicts *t2* in dementia contexts. As Dresser remarks, "few individuals know [beforehand] how they would experience the relevant burdens and benefits in their incapacitated state"[27] (p. 3). Thus, while *t1* can provide important evidence, it cannot by itself settle what the person now prefers. At best, it offers partial and defeasible guidance. Overall, authenticity in advanced dementia cannot, in general, be settled either by counterfactual endorsement alone or by treating *t1* preferences as automatically authoritative, since both routes leave substantial room for underdetermination in precisely the cases that HCPs struggle with. The practical upshot is that clinical contexts require defeasible, usable indicators for when it is reasonable to treat an expressed preference as authentic and morally weighty, and when less so. The section six proposes such indicators.

Situating the View in the Literature: Two Routes

Having set out our proposed framework and considered two natural objections to it, we can now situate the view in relation to existing debates. A number of authors agree that some form of decisional authority may outlive full decision-making capacity, and they offer different ways of conceptualising the resulting normative remainder. Our account shares this starting point that the lack of cognitive competence does not straightforwardly erase a person's status as someone whose preferences continue to matter. However, our account develops the remainder in a distinctive way by grounding it in a thin, historically anchored notion of authenticity. This section briefly maps two main routes in the literature and explains how our appeal to Christman is meant to build on their insights while addressing their limitations.

One influential route emphasizes *valuing* rather than competence. Jaworska, in particular, argues that what underwrites morally significant agency is not sophisticated decision-making ability linked to autonomy or a "sense of one's life as a whole," but the capacity to value[28]. On her account, although many people with dementia are unable to form new values, they retain evaluative attitudes that are meaningfully distinct from mere impulses, and it is this capacity to

value that grounds their continuing moral standing and a form of residual authority.

At the same time, Jaworska explicitly acknowledges that authenticity matters morally: pathology-generated values would carry "not nearly as much moral weight" as values that are authentic, understood as values "anchored in the mental history of the living person." Yet she then largely brackets this "authenticity complication," noting that Alzheimer's patients' values are "typically authentic in this sense"[28] (p. 124). Our approach differs in two respects. First, we do not treat authenticity as something that can be assumed to hold "typically" in dementia contexts. Instead, we treat cases of possible distortion as central, precisely because HCPs often face situations where it is unclear whether an expressed preference is authentic or primarily disorder-driven. Second, we do not construe authenticity as requiring anchorability in a pre-dementia standpoint in a way that would automatically privilege t1 over t2: on our thin, procedure-externalist account, the question is whether the *present* preference is authentic, and we therefore aim to provide defeasible grounds for presuming authenticity in some cases rather than simply assuming it "typically" holds.

A second route focuses on authenticity as distinct from mere decisional competence. In line with our approach, Brudney and Lantos suggest that HCPs' discomfort with some competence-confirmed refusals of treatment is best understood as an authenticity judgment rather than as a best-interests judgment[8]. However, unlike our account, they develop a thick, diachronic notion of authenticity. While they think that agency refers to the capacity to choose on the basis of reasons, they hold that authenticity is fidelity to one's enduring values and practical identity, which is something that "must be exercised over time, sometimes over a lifetime"[8] (p. 121). Authenticity concerns whether a choice is understood as a sustained achievement of "making a life" that coheres with a relatively stable pattern of commitments and values over time.

However, in the context of advanced dementia, there are reasons to doubt whether this is appropriate. First, it is vulnerable to our reply *Past-Self Privileging Objection*, which argued that authenticity should not be construed as something that is only anchorable in a pre-dementia history. Second, where diachronic narrative coherence is often disrupted and

inaccessible, this standard risks being too demanding to guide judgments about the moral weight of present preferences. In reply, Brudney and Lantos could argue that authenticity need only involve alignment with a person's enduring values and practical identity, not the person's current access to them. However, alignment alone is not sufficient since it may be merely coincidental: a disease-produced preference may "fit" a past outlook by accident or through prompting. To avoid this, the relationship between a particular present preference and enduring values and practical identity that makes this preference authentic must be more than superficial fit: the former must be appropriately grounded in and arising from the latter. This, however, plausibly requires some present access to the relevant identity-defining commitments, and since advanced dementia often disrupts precisely this kind of access, thick authenticity becomes too demanding for dementia cases.

Taken together, the two prominent routes in the literature capture important parts of the phenomenon our workshop cases bring into view. However, while Jaworska largely brackets the "authenticity complication", and Brudney and Lantos' thick conception risks setting the bar too high, our account is better suited to that setting. It preserves the intuition that some present preferences can remain authentic, provides a principled way to articulate authenticity, and supports the practical step that next section sets out to develop: identifying defeasible indicators that can serve as reasonable evidence of authenticity in everyday clinical practice.

Conditions for Adopting Authentic Preferences

Before turning to these indicators, it is important to be explicit about a limitation of our approach. In many cases, it will not be possible to determine with certainty whether the preferences expressed by a person with advanced dementia are authentic. We thus largely endorse Holm's conclusion that a substantial grey zone will remain: no single criterion can determine, in all cases, whether a preference should be respected or overridden, and caregivers cannot be spared the need for contextual, interpretive judgment[29]. Our claim is therefore defeasible: there are identifiable indicators which, when present, provide sufficient justification for presuming authenticity

despite the absence of autonomy. The claim is that when people with advanced dementia exhibit three indicators of authenticity, this provides *pro tanto* grounds for treating their expressed preferences as authentic. The framework of indicators comprises: (1) consistency in preferences, (2) minimal comprehension, and (3) the absence of salient signs of external influence.

The first indicator is consistency in preferences, which we understand as a single, clinically usable check with two distinguishable components: minimal prioritisation and temporal stability. The first is not the stronger claim in parts of the autonomy literature that autonomy requires that a person's beliefs and desires "do not contain manifest inconsistencies" [30] (p. 278). Rather, following Oshana's emphasis on ranking what matters and Frankfurt's notion of the wanton, what is especially salient for our purposes is whether the person has any capacity to treat some preferences as having priority over others [31]. Where this capacity is entirely absent, the person is left unable to resolve clashes and thus lacks direction. Importantly, we do not treat this kind of prioritisation as requiring the demanding form of critical self-reflection often associated with autonomy, since that would set the bar too high for our purposes. Even agents who plausibly lack such reflective capacities (e.g., young children) can nonetheless display rudimentary prioritisation and relatively stable patterns of what matters to them. At the same time, prioritisation without temporal stability would provide little evidence of an even minimally settled standpoint. The relevant notion of stability is not behavioural repetition, which may reflect little more than motor habit, but stability of preference across time and across at least some variation in context. In practice, a pattern of refusals or endorsements that recurs over days or weeks, and that appears in different but related situations, provides stronger evidence of a stable evaluative stance than a single incident that can also be explained by other factors such as agitation, confusion, or environmental triggers.

The second indicator is a minimal degree of comprehension. This also echoes familiar themes in autonomy theory. For example, Christman characterises adequate understanding as the ability to "realistically imagine choosing otherwise were she in a position to value sincerely that alternative position" (p. 280) [30]. However, for our purposes it is important

not to set the bar too high. Understanding is not an all-or-nothing capacity, but a domain-specific phenomenon we refer to as *spheric* understanding. One may competently evaluate options in one sphere (e.g., choosing between coffee and tea) while being unable to do so in another (e.g., choosing a new hobby). Understanding may be retained in some spheres but not others, and it may persist for simpler matters while failing for more complex ones. The spherical model of understanding thus frames dementias as conditions of progressive narrowing, whereby the capacity to grasp abstract and complex matters erodes, leaving comprehension increasingly confined to simpler, concrete domains [28, 32, 33]. Moreover, a minimal degree of comprehension requires only that the individual grasp the immediate consequences of their choice, not that they possess the broader capacities necessary to assess its longer-term implications. For example, consider a patient who repeatedly expresses a preference to be alone and to avoid social interaction. On our account, this preference satisfies the minimal-comprehension indicator insofar as the patient shows a rudimentary understanding of what it involves, namely, that avoiding others will result in being left alone and may, over time, lead to isolation. Still, she may lack an understanding of the potential emotional or psychological harm that prolonged isolation might entail, or the possibility that she could derive satisfaction from future social interactions.

The third indicator is the absence of salient compelling medical or non-medical external influence that distorts or overrides preferences. This is a familiar requirement across most plausible accounts of autonomy, which typically hold that self-determination is undermined when apparent choices are driven by coercion, manipulation, or pathology rather than by the agent's own evaluative standpoint. Accordingly, authenticity is compromised when what looks like a choice is better explained as the product of delusion, severe disturbance, or other reflection-distorting factors than as an expression of the person's preferences. Consider a person with advanced Lewy Body dementia who, because of dementia induced paranoia, refuses food, being convinced that her caregivers have poisoned it. Such a refusal is not authentic: it arises from delusional pathology rather than a genuine evaluative outlook. Equally, dysphagia may prevent eating despite a genuine desire. Some kinds of undue external influence may be more easily

detectable to caregivers than others. For instance, it may be easier to detect dysphagia or delirium than, e.g. depression in a person with advanced dementia.

Taken together, these three indicators provide a practical framework for evaluating authenticity in dementia. Accordingly, caregivers are justified in treating an expressed preference as authentic only insofar as, to the best of their ability, they can establish that it is stable over time, minimally understood by the person, and not plausibly driven by distorting external factors. However, although authenticity can ground a *prima facie* reason to respect certain preferences of people with advanced dementia, this does not mean that such preferences cannot be justifiably overridden by countervailing considerations. For example, where a person with dementia expresses multiple, potentially conflicting preferences, the prioritisation component of the consistency indicator is relevant: a pervasive inability to resolve conflicts weakens the justification for treating any single expression as action-guiding on authenticity grounds. And even when a preference is well-supported by our indicators, differences in what caregivers ought to do need not be explained by "more" or "less" authenticity. Rather, what varies is the practical authority it is reasonable to grant the preference, which depends (i) on how strongly it is supported by the three defeasible indicators and (ii) on the practical stakes and potential harms of compliance: high-stakes decisions require stronger evidential support and are more likely to be overridden by countervailing considerations.

Implications and Real-life Cases

Let us now return to the real-life cases that motivated this discussion and explore the implications of the upshot so far: even though people with advanced dementia are typically unable to engage in higher-order critical reflection on their preferences, it does not follow that they are incapable of holding authentic preferences, nor that caregivers are unjustified in presuming authenticity when preferences are expressed through behaviour displaying the relevant indicative features. One straightforward implication is that the account defended here helps make sense of the pattern observed among the HCPs in the workshop. Of course, without further investigation we cannot know whether the HCPs were *in fact*

guided by these considerations. But to the extent that their judgments track the indicators of authenticity we have identified, they would be justified on our account.

The workshop cases concerned preferences that appeared to conflict with a resident's prior identity and commitments: (1) a lifelong vegetarian with dementia who suddenly craved meat, (2) a Muslim resident who requested pork, and (3) residents who refused food and water. The HCPs tended to treat the vegetarian case as negotiable: many initially hesitated, but often accommodated the new desire. However, in the pork case, they were more reluctant, and in the refusal-of-nutrition case they were most willing to respect the expressed preference. Our account helps make sense of this pattern by showing how the moral weight of present preferences can depend on whether they plausibly satisfy the indicators of authenticity. For the purposes of the discussion, we will assume that the relevant preference is stable over time, since a one-off expression of preference would be a non-starter: it provides too little evidence to justify treating the preference as authenticity-grounded in the first place.

In the vegetarian case, it is plausible that the relevant criteria are perceived as met. Minimal comprehension is comparatively straightforward to ascribe: the resident can plausibly recognise what "meat" is (as opposed to other foods), express a stable choice for it, and display signs of uptake (e.g., accepting it when offered, rejecting alternatives). Likewise, the absence of undue external influence is plausible because there is typically no clear strong norm that would pressure compliance to consuming meat. In combination, these features make it reasonable, on our framework, to treat the preference as a candidate for authenticity. Further, because abstaining from meat is often treated as a commitment whose violation does not necessarily signal a deep rupture in practical identity, accommodating the request can be more like recognising a permissible change in present preference.

In the pork case, it is plausible that one or more of the indicators is not satisfied, which would explain the HCPs' greater reluctance even if we assume equal stability over time. First, minimal comprehension may be less clear. The resident might understand concretely what pork is and that they are requesting it, yet still fail to grasp what is normatively salient about

the choice for them: for example, that pork was previously treated as forbidden or identity-significant within their religious outlook. If caregivers suspect that this evaluative significance is no longer accessible to the resident in their present cognitive state, then the second indicator is not straightforwardly met, even if the preference is persistent. Second, the absence of undue external influence may also be less clear. Even without overt coercion, caregivers may reasonably worry that the request reflects heightened susceptibility to background social norms in the surrounding environment, pressures that a cognitively impaired person may be less able to resist. On that interpretation, the request can look less like an expression of the person's evaluative standpoint and more like acquiescence to an external normative environment.

Taken together, these considerations make it somewhat understandable that HCPs were more hesitant in the pork case: even if the preference is stable, doubts about minimal comprehension and freedom from external pressures weaken the justification for treating the preference as authentically the person's own. With that said, we cannot rule out that the absence of consensus in the pork case was partly due to the fact that many participants were themselves Muslim HCPs, so that their own religious commitments may have influenced how they assessed the appropriateness of complying with the request. Some evidence that this kind of effect *could* be implicated comes from empirical work on "true self" judgments, which suggests that judgments whether someone's actions or preferences are authentic can be influenced by the evaluator's own value-laden outlook (Newman et al., 2014). Applied to the pork case, this raises the possibility that some Muslim HCPs may have been more inclined to regard the resident's request for pork as *not* reflecting the resident's authentic preferences (and thus as carrying less normative weight), in part because it conflicts with the HCPs' own religious commitments. We emphasize that we do not know whether this in fact occurred in the workshop. But we flag it as a potential confound, a sort of bias our content-neutral indicators are meant to help guard against.

Finally, can our account make sense of the HCPs' relative willingness to respect refusals of food and water? Assuming, for simplicity, that the refusal is stable over time, our account makes the HCPs' relative willingness to respect it intelligible. First, it can

be plausible that the patient retains minimal comprehension of cause and effect, that continued refusal of food and fluid will lead to death, even if higher-order reflections about life are no longer accessible. Although such decisions would ordinarily be bound up with complex, higher-order considerations (e.g., religious commitments, family obligations) that are often inaccessible in advanced dementia, it does not follow that all understanding disappears. Second, the preference can also plausibly be taken to be free from undue influence, that is, not best explained by distorting factors like delusion (e.g., paranoia about poisoning), unmanaged pain, or coercion. Thus, the refusal can reasonably be treated as a candidate for thin authenticity and thus as carrying *prima facie* moral weight. That said, there may also be an additional factor at work: because hunger and thirst are among the most basic human drives, if a person repeatedly forgoes such basic needs, HCPs may be inclined to see this as signaling something more than a fleeting impulse, and therefore as particularly likely to be authentic in the relevant sense.

Overall, the account defended here helps make intelligible the pattern observed among the HCPs in the workshop. Of course, without further investigation we cannot know whether the HCPs were *in fact* guided by these considerations and the three authenticity indicators. Assessing that would require evidence not only about the reasons they offer, but also about their decision-process in practice (e.g., through vignette studies, qualitative interviews). Still, insofar as their judgments track the indicators of authenticity we have identified, they would be justified on our account.

Further Implications

Much of the discussion in contemporary debates is structured around a choice between respecting autonomy (often operationalised through decision-making capacity) and promoting welfare through best-interest reasoning. Our account complicates that framing by identifying a distinct source of normative weight: where a preference is plausibly authentic, it can merit respect even in the absence of full autonomy and even when welfare considerations do not straightforwardly settle the case. In this way, the proposal both offers one way to explain why HCPs often hesitate to treat

incapacity as automatically eliminating decisional authority and provides a principled basis for that hesitation, while remaining compatible with the thought that such reasons are defeasible and may be overridden in cases of serious harm or compelling countervailing considerations.

To clarify what, exactly, this contribution amounts to in relation to existing debates, it is helpful to distinguish three respects in which our account both builds on and departs from prominent approaches in the literature. It departs from (1) views that ground the moral weight of dementia patients' preferences in considerations of personhood or liberty, (2) welfare approaches that justify compliance primarily by appeal to well-being, and (3) approaches that already invoke authenticity.

First, our approach differs from other strategies for grounding obligations to respect the preferences of people with advanced dementia. Wasserman and Navin, for instance, argue that even when patients lack decision-making capacity, their expressed preferences retain moral weight because respecting them affirms personhood and liberty beyond considerations of autonomy or best interests[13]. By contrast, our argument does not rely on external appeals to liberty or personhood. Instead, we maintain that HCPs have an obligation, albeit a weak one, to respect such preferences simply because, like all people, individuals with dementia are owed respect for their genuine preferences.

Second, from both Benthamite instrumental and Millian intrinsic perspectives, the case for respecting preferences is closely tied to welfare considerations, which can generate plausible reasons for giving substantial weight to these preferences in dementia care[34]. On the instrumental view, respecting autonomy tends to promote welfare because individuals are generally best placed to judge what benefits or harms them, while the intrinsic view adds that exercising one's will constitutes a constituent good of a flourishing life. On either approach, respecting the preferences of people with advanced dementia may promote welfare both by satisfying their expressed desires and by allowing some exercise of volition. However, while some of the reasons for adhering to the relevant preferences may indeed derive from their potential to promote welfare, an essential part of the moral rationale is independent of welfare. Because authentic preferences are expressions of the person's

evaluative standpoint, they can generate *prima facie* reasons of respect even when complying does not clearly promote welfare.

Third, and finally, our account also contributes to that strand of the literature which treats authenticity as morally relevant in dementia ethics[27, 35], but which often leaves the concept under-specified or tacitly equates it with simple continuity with past identity. By embedding authenticity within an autonomy framework and operationalising it through defeasible indicators, we clarify what authenticity amounts to in this context and how it can plausibly ground *prima facie* reasons for respect even when full autonomy is absent.

Concluding Remarks

This paper set out to clarify whether, and in what sense, the expressed preferences of individuals with advanced dementia can retain moral significance despite the loss of full autonomy. Motivated by clinical observations that HCPs often treat some preferences as still expressive of the person's evaluative outlook, we develop an account that both vindicates these intuitions and renders them operational through defeasible indicators. To do so, we defend a thin, procedure-externalist conception of authenticity that can, in principle, persist even when higher-order reflective capacities have collapsed. We argued that authenticity remains conceptually available even in advanced dementia, and that it can ground a *prima facie* normative reason for HCPs to respect certain preferences, independent of welfare and not reducible to precedent autonomy as expressed in advance directives. Moreover, to make this claim action-guiding, we proposed three defeasible indicators (consistency of preferences, minimal comprehension, and the absence of undue influence) that can serve as reasonable evidence of authenticity in everyday clinical practice.

Of course, this account is necessarily limited in scope, and many further questions remain. First, the three indicators could be refined, and future research should assess their reliability in clinical decision-making. Second, it remains an open question whether authenticity should be understood as partly relational, shaped, at least in part, by the evaluative contributions of family members and caregivers. Third, we do not claim to have shown that the HCPs in our workshop

were in fact guided by authenticity considerations. Whether clinicians rely on these indicators (explicitly or implicitly) is an empirical question that would require further study, for example through interviews and vignette-based designs that systematically vary stability, comprehension, and undue influence. Fourth, and on a larger philosophical scale, the possibility of authentic preferences in the absence of autonomy highlights the need for autonomy theories to accommodate forms of agency that are historically shaped, environmentally scaffolded, and often enacted without reflective endorsement. Future work will need to clarify how these forms of agency fit within existing taxonomies of autonomy and responsibility.

Nevertheless, despite these and many other open issues, the central point remains: respecting people with dementia requires more than safeguarding their welfare or enforcing their past autonomous wishes. It requires recognising the residual evaluative standpoint they still occupy, however fragile or intermittent. Authenticity, in its thin but morally important sense, provides a way of acknowledging and responding to that standpoint.

Author Contributions S.L. wrote the first draft of the manuscript, and S.V. commented on earlier versions and carried out most revisions after reviewer feedback. All authors read and approved the final manuscript.

Funding Open access funding provided by University of Southern Denmark.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Lauridsen, S., F. Schou-Juul, A. P. Folker, P. Simonsen, M.-E. Phil, and S. S. Skov. 2023. Developing the CARE intervention to enhance ethical self-efficacy in dementia care through the use of literary texts. *BMC Medical Ethics* 24 (1): 45.
2. Lauridsen, S., S.S. Skov, L.-M. Thorup, M.-E. Berg, A.P. Folker, P. Simonsen, and F. Schou-Juul. 2025. Enhancing ethical self-efficacy among dementia caregiver staff: A mixed-methods feasibility study of the CARE programme. *International Journal of Nursing Studies Advances* 8: 100294.
3. Rawls, J. 1980. Kantian constructivism in moral theory. *The Journal of Philosophy* 77 (9): 515–572.
4. Feinberg J. 1987. *The moral limits of the criminal law*: Volume 1: harm to others. Oxford University Press, 1989.
5. Friedman, D. 2013. Public health regulation and the limits of paternalism. *Connecticut Law Review* 46: 1687.
6. Grisso, T., and P. S. Appelbaum. 1995. MacArthur treatment competence study. *Journal of the American Psychiatric Nurses Association* 1 (4): 125–127.
7. Grisso, T., P. S. Appelbaum, and C. Hill-Fotouhi. 1997. The MacCAT-T: A clinical tool to assess patients' capacities to make treatment decisions. *Psychiatric Services (Washington, D.C.)* 48 (11): 1415–1419.
8. Brudney, D., and J. Lantos. 2011. Agency and authenticity: Which value grounds patient choice? *Theoretical Medicine and Bioethics* 32:217–227.
9. Pennington, C., K. Davey, R. ter Meulen, E. Coulthard, and P. G. Kehoe. 2018. Tools for testing decision-making capacity in dementia. *Age and Ageing* 47 (6): 778–784.
10. Poppe, C., B. S. Elger, T. Wangmo, and M. Trachsel. 2020. Evaluation of decision-making capacity in patients with dementia: Challenges and recommendations from a secondary analysis of qualitative interviews. *BMC Medical Ethics* 21 (1): 55.
11. Christman J. 2005. Procedural autonomy and liberal legitimacy. *Personal autonomy: New essays on personal autonomy and its role in contemporary moral philosophy* Taylor, J. S. (Ed.). 277–298. Cambridge University Press.
12. Noggle R. 2005. Autonomy and the paradox of self-creation: Infinite regresses, finite selves, and the limits of authenticity. Taylor, J. S. (Ed.). 87–109. Cambridge University Press.
13. Wasserman, J. A., and M. C. Navin. 2018. Capacity for preferences: Respecting patients with compromised decision-making. *Hastings Center Report* 48 (3): 31–39.
14. Jongsma, K. R., M. A. G. Sprangers, and S. van de Vathorst. 2016. The implausibility of response shifts in dementia patients. *Journal of Medical Ethics* 42 (9): 597–600.
15. Kinch S., Schou-Juul F., Skov S.S., Kongsholm N.C.H., and Lauridsen S. 2024 The use of advance directives for autonomy in dementia care: A scoping meta-review and thematic synthesis. *Archives of Gerontology and Geriatrics* 105498.
16. Quante, M. 2011. In defence of personal autonomy. *Journal of Medical Ethics* 37 (10): 597–600.

17. Taylor, C. 1992. *The ethics of authenticity*. Harvard University Press.
18. Varga, S. 2013. *Authenticity as an ethical ideal*. Routledge.
19. Darwall, S. 2009. *The second-person standpoint: Morality, respect, and accountability*. Harvard University Press.
20. Earp, B. D., J. Demaree-Cotton, M. Dunn, V. Dranseika, J. A. Everett, A. Feltz, G. Geller, I. R. Hannikainen, L. A. Jansen, and J. Knobe. 2020. Experimental philosophical bioethics. *AJOB Empirical Bioethics* 11 (1): 30–33.
21. Toomey, J., J. Lewis, I. R. Hannikainen, and B. D. Earp. 2024. Advance medical decision-making differs across first-and third-person perspectives. *AJOB Empirical Bioethics* 15 (4): 237–245.
22. Earp B.D., Hannikainen I.R., Dale S., Latham S. 2023. Experimental philosophical bioethics, advance directives, and the true self in dementia. *Advances in experimental philosophy of medicine Bloomsbury* 259–289.
23. Schechtman, M. 1996. *The Constitution of Selves*. Ithaca, NY: Cornell University.
24. Boorse, C. 1976. What a theory of mental health should be. *Journal for the Theory of Social Behaviour*. 44 (4): 542–573.
25. Stirland, L. E., B. A. Ayele, C. Correa-Lopera, and V. E. Sturm. 2023. Authenticity and brain health: A values-based perspective and cultural education approach. *Frontiers in Neurology* 14 : 1206142.
26. Wilkins, J. M. 2021. Reconsidering gold standards for surrogate decision making for people with dementia. *Psychiatric Clinics of North America* 44 (4): 641–647.
27. Dresser, R. 2009. Substituting authenticity for autonomy. *Hastings Center Report* 39 (2): 3.
28. Jaworska, A. 1999. Respecting the margins of agency: Alzheimer's patients and the capacity to value. *Philosophy & Public Affairs* 28 (2): 105–138.
29. Holm, S. 2001. Autonomy, authenticity, or best interest: Everyday decision-making and persons with dementia. *Medicine, Health Care and Philosophy* 4 (2): 153–159.
30. Christman, J. 2007. Autonomy, history, and the subject of justice. *Social Theory and Practice* 33 (1): 1–26.
31. Oshana, M. 2007. Autonomy and the question of authenticity. *Social Theory and Practice* 33 (3): 411–429.
32. Mancano, M., and C. Papagno. 2023. Concrete and abstract concepts in primary progressive aphasia and Alzheimer's disease: A scoping review. *Brain Sciences* 13 (5): 765.
33. Joubert, S., G. T. Vallet, M. Montembeault, M. Boukadi, M. A. Wilson, R. J. Laforce, I. Rouleau, and S. M. Brambati. 2017. Comprehension of concrete and abstract words in semantic variant primary progressive aphasia and Alzheimer's disease: A behavioral and neuroimaging study. *Brain and Language* 170:93–102.
34. Mill, J. S. 1859. *On Liberty*. London: J.W. Parker.
35. Carter, M. 2022. Advance directives: The principle of determining authenticity. *Hastings Center Report* 52 (1): 32–41.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.