DOSTOEVSKY'S GRAND INQUISITOR: THE RELIEF OF OBEDIENCE AND THE LIMITS OF RESPONSIBILITY

Maurília Gabriela da Silva Dietrich

Centro Universitário Leonardo da Vinci – UNIASSELVI. Bachelor of Laws; postgraduate student in Religious Studies and Philosophy Applied to Religion and Education at Faculdade Líbano; undergraduate student in Theology and in History (teaching degree) at Faculdade Estácio de Sá · ORCID 0009-0003-2756-6848 · mauriliagabriela@outlook.com

Pablo Dietrich

Fundação Universidade Regional de Blumenau – FURB. Bachelor of Laws; postgraduate student in Religious Studies and in Philosophy and Sociology at Faculdade Líbano; undergraduate student in Theology and Philosophy at Faculdade Estácio de Sá · ORCID 0009-0003-5447-0550 · pablodietrichh@hotmail.com

Authors' translation. English translation, prepared by the authors, of: DIETRICH, Maurília Gabriela da Silva; DIETRICH, Pablo. O grande inquisidor de Dostoiévski: o alívio da obediência e os limites da responsabilidade. *Revista Nexus Interdisciplinar*, Natal/RN, v. 1, n. 3, p. 764–785, Jul./Sept. 2026. ISSN 3086-4402. DOI: 10.47538/RNI-2026.V1N3-43. The Portuguese text published by the journal is the version of record; this translation is distributed under the CC BY 4.0 licence with attribution to the original publication. Quotations from works consulted in Portuguese translation or in other languages, as well as from the authors' own works, are rendered into English by the authors.

ABSTRACT

The literature on obedience treats as a single block, under headings such as escape from freedom, two effects that this article separates. Decisional relief is the exemption from the work of deliberating; imputational relief is the exemption from understanding oneself as the normative source of the act. It is argued that these are distinct states, and the proof is the difference of their defeaters, that is, of the operations capable of dissolving them: the first dissolves when the decision is returned to the agent, even if he remains convinced that the responsibility belongs to someone else; the second dissolves when he is shown that the transfer did not occur, without any part of the deliberation being returned to him. It follows from the distinction that only the second contains an assessable normative claim, which transfers it from moral psychology to moral philosophy. Once that claim is assessed, three results are established. First, it is false where the incapacity at stake is moral, in Edwards's sense, and true only to the extent of the epistemic conditions that the agent could not have altered. Second, there is a structurally non-transferable residue, which lies not in the decision but in the judgement by which the authority was recognized, since delegating it would require exercising it; it is argued, on independent grounds and in dialogue with the device of tracing, that obligation derives from the position occupied and not from present capacity. Third, once liability to consequences, transferable by convention, is distinguished from blameworthiness and from answerability, it turns out that responsibility is not a conserved quantity: a third party's willingness to answer for the act adds what he has to answer for without subtracting what the agent has. The episode of the Grand Inquisitor, in Book V of *The Brothers Karamazov*, is used as an exemplary case, since in it the offer of transfer is explicit and accompanied by a declared justification.

Keywords: moral responsibility; imputation; obedience; blameworthiness; Dostoevsky.

INTRODUCTION

Whoever obeys is spared from deciding. The observation is old and true, and it underpins a good part of what has been written about the attraction exerted by authority: deliberating is burdensome, uncertainty is painful, and the possibility of erring irreparably is permanent. If the problem were only this, obedience would be an arrangement of cognitive comfort, and its moral assessment would be exhausted by asking whether the decision taken by someone else was sound.

There is, however, a second effect, which the literature records without separating it from the first. Besides being spared from deciding, the agent who obeys receives a moral description of his own conduct under which he can recognize himself without scandal: it was my duty, the order was legitimate, the decision was not mine, I obeyed a higher norm. It is not a matter of ceasing to be the cause of the act, for he knows perfectly well that it was his hands; it is a matter of ceasing to understand himself as the instance from which the act receives its moral qualification. The first effect will be called decisional relief and the second imputational relief.

This article argues that they are two states, and not one state accompanied by a belief about it, and that the distinction alters the assessment of both. Decisional relief is a primarily psychological and organizational phenomenon, verifiable without anything being decided about imputation. Imputational relief is constituted by a normative claim, for whoever experiences it implicitly asserts something about who answers for the act; and a normative claim can be true or false. Asking whether it is true ceases to be a question of moral psychology and becomes a question of moral philosophy.

The argument unfolds in four steps. First, the duality is established through the difference of the defeaters, that is, by showing that each of the two reliefs is dissolved by an operation that does not dissolve the other. Second, the claim contained in the second relief is assessed, and it is shown to be false where the incapacity at stake is moral, and true only to the extent of the epistemic conditions that the agent could not have altered. Third, a structurally non-transferable residue is identified, located not in the decision but in the judgement of recognition of the authority, and it is argued, on independent grounds, that obligation derives from the position occupied and not from present capacity. Fourth, once three notions that ordinary language conflates have been distinguished, it is shown that responsibility is not a conserved quantity, and that the offer to take on another’s burden has the grammar of a substitutive imputation without the structure that would make it intelligible.

Two declarations of procedure must precede the argument. The first is that it is reconstructive and analytical-conceptual; the second, that it is neutral, and neutrality is not presumed: it is stated. No argument is made for or against any creed, nor is the truth of the religious reasons invoked by the case under analysis examined; what is examined is what an offer of transfer produces and what it does not produce. Two conceptual tools will be mobilized, and it is fitting to announce them: the distinction, elaborated by Jonathan Edwards, between natural inability and moral inability, with the consequence that only the first excuses (Edwards, 1957,

I.4), argued by Edwards himself from the concept of inability and not by doctrinal authority; and the thesis that obligation derives from the position occupied, which will not be presupposed, but rather established further on by an argument of its own, in its purely formal form and without identifying any position as fundamental.

THE CASE AND THE SUPPLEMENTARY OFFER

The case examined is the episode of the Grand Inquisitor, in Book V of *The Brothers Karamazov*, and the choice is justified by a precise reason: in it the offer of transfer is explicit, accompanied by a declared justification and formulated by a character who is not presented as cynical. The episode functions as an exemplary case, and not as an object of literary interpretation; the apparatus of literary criticism is reduced to the minimum necessary to ensure the correct attribution of the voices.

No assessment of the episode can dispense with its prior reconstruction, and the first care that this requires concerns attribution. The poem is not an essay by Dostoevsky: Ivan Karamazov claims to have composed it in his head about a year earlier and narrates it to Alyosha in a tavern, ironically and admitting its literary crudeness. At least four instances are therefore to be distinguished — the author, the narrator of the novel, Ivan, and the Inquisitor — and the polyphonic structure of the novel, as Bakhtin (2010) described it, is the reason why none of them can be taken as the author's statement. The episode is treated here as an inquiry in philosophical anthropology, in the key opened by Sandoz (1964), and not as a piece of doctrine or an object of literary exegesis.

The relevant narrative elements can be gathered economically. Christ returns silently to Seville, in the sixteenth century, on the day after a great auto-da-fé, and is recognized by the people; the Grand Inquisitor has him arrested and visits him at night in his cell, where he delivers a long monologue. He accuses him of having overestimated human strength and of having refused, in the desert, the three means capable of definitively conquering the conscience of men — miracle, mystery, and authority —; he asserts that the Church has corrected that work; he declares that he acts for the benefit of a weak humanity, to which he will give bread, order, and the peace of not needing to choose; and he announces that Christ will be burned the next day. Christ does not reply discursively: he approaches and kisses him. The old man opens the door and sends him away, and remains faithful to his idea (Dostoiévski, 2008, Book V, ch. 5).

One observation precedes any theoretical apparatus, and it concerns the character's vocabulary. The Inquisitor does not speak the language of tyranny, but that of care: he describes himself as the one who suffers in the place of others, who carries the burden of knowledge, and who accepts the curse so that the multitude may remain happy. To treat this discourse as hypocrisy is to replace a difficult problem with an easy adversary; what the text demands is that one take seriously the possibility of a sincerely benevolent domination, and it is with this possibility, and not with cynicism, that the article is concerned.

The established reading of the episode opposes freedom and security: the old cardinal accuses Christ of having demanded of men a freedom they are unable to bear, and the Church he leads would have corrected the error by assuming the burden of choice. The reading is sound and captures something true, but it describes the demand side and remains vague as to the supply side. The Inquisitor does not say only that he will decide for others. He says that he will allow men to sin and will take their sin upon himself, that they will die peacefully knowing that someone will answer in their place, and that obedience will be presented to them not as resignation but as faithfulness. It is this supplementary offer, and not the first, that the present article takes as its object.

It is worth recording, without developing it, the framework within which the offer becomes possible, for only what the argument requires is drawn from it. An authority may be legitimate even if it pre-empts a considerable part of the agent's deliberation, and Joseph Raz's service conception explains why: it helps him to conform better to the reasons that already apply to him (Raz, 1986; 2006). The arrangement examined here exceeds that framework in two respects — the scope of the competence claimed, which is not delimited by domains, and the removal of the matters in which deciding for oneself has a weight of its own. What results from this excess is described precisely by the theory of noumenal power: someone exercises power over another when he is able to occupy that other's space of justifications, so that the latter acts for reasons that the former has supplied, and the most stable form of this exercise is the one in which submission presents itself to the subject as the conclusion of his own moral reasoning (Forst, 2015).

This framework establishes what the authority has done to the agent; it does not establish what the agent still has to answer for. It is one thing to show that someone's space of reasons has been occupied, another to determine whether, and to what extent, the occupation alters imputation. It is this second question that the rest of the article addresses, and it requires, before anything else, that what the occupation spares the agent at the moment of deciding be separated from what it offers him at the moment of answering.

DECISIONAL RELIEF AND IMPUTATIONAL RELIEF

We now reach the propositional core of the article. The thesis is that the arrangement described produces two distinct effects, usually treated as a single block under the generic heading of escape from freedom (Fromm, 1941), and whose separation alters the assessment of both.

Decisional relief consists in the exemption from the work of deliberating: the subject does not need to interpret the situation alone, weigh conflicting reasons, bear uncertainty, or choose between morally difficult alternatives, for the authority determines the conduct and the cognitive and emotional cost of the determination disappears. Imputational relief consists in the exemption from understanding oneself as the normative source of the act: the authority supplies not only the conduct, but the moral description under which the conduct can be recognized — it was my duty,

the order was legitimate, the decision was not mine, I obeyed a higher norm, my action served a greater good.

The distinction is finer than it seems, and it should be stated precisely. The agent who experiences imputational relief does not, in general, claim to disappear as the cause of the act, for he knows perfectly well that it was his hands. What he ceases to recognize is that he was the instance from which the act received its moral qualification. He remains the material author and ceases to understand himself as the normative author.

The two reliefs can come apart, and the possibility of dissociation is the proof that they are indeed two. There is decisional relief without imputational relief: the patient who follows medical advice is spared technical deliberation and continues to recognize himself fully as the one who chose to be treated and who will answer for that choice. And there is imputational relief without decisional relief: the agent who deliberates at length, chooses between alternatives, calculates consequences, and nevertheless describes the result as the fulfilment of a design that does not belong to him. It is in this second case, and not in the first, that the most characteristic feature of the Inquisitor's project is found, for deliberation may even intensify; what dissolves is normative authorship.

From this results the circuit that the episode dramatizes, and which it is worth tracing in full. The authority presents itself as the interpreter of higher reasons; the subject delegates part of his practical judgement; the institutional directive comes to be perceived as a duty; obedience produces decisional relief; the justification supplied by the authority produces imputational relief; and the double relief reinforces the recognition of the very authority that produced it. The circuit presupposes neither conscious agreement nor deliberate manipulation, and the poem constructs the Inquisitor in exactly this way: authorities and subjects may sincerely believe in the reasons invoked. Sincerity explains the internal force of a justification; it does not demonstrate its validity.

Moral psychology offers limited but pertinent support to the second relief. Bandura (1999) described a set of mechanisms of moral disengagement by which harmful conduct can be reconstructed as benign or meritorious without any alteration of the agent's self-image, among them moral justification, euphemistic labelling, advantageous comparison, and the diffusion and displacement of responsibility. Two points of this description are of interest.

The first is that the displacement of responsibility is not an operation the agent performs alone, for it requires an authority willing to accept the burden, or at least to present itself as one who accepts it — and the Inquisitor does precisely this when he declares that he will take the sin of men upon himself, so that what appears in the psychological literature as an individual mechanism finds in the poem its explicit institutional counterpart. The second is that the mechanisms can operate before the conduct, and not only after it, for nothing in the description requires the reconstruction to be retrospective; a justification already available would dispense with the subsequent work of reconstruction and with the discomfort that would motivate it. It is not asserted here, since there is no basis for asserting it, that the anticipatory form is empirically

more effective than the retrospective one; it is asserted only that the authority described by Ivan operates in the first form. The authority described by Ivan does not offer absolution afterwards; it offers the moral description beforehand.

The scope of this contribution must be rigorously delimited, and the delimitation is the same that the literature on moral rationalization already imposes on itself (Tsang, 2002). Bandura provides the explanation of why imputational relief works psychologically, and he provides it with considerable empirical precision. What he does not provide is any conclusion as to whether the relieved agent is, in fact, less responsible; from the efficacy of a mechanism no normative assessment whatsoever follows, and to treat moral psychology as the foundation of the philosophical thesis would be to confuse causal explanation with justification. What can be drawn from it is more modest and yet relevant: when the feeling of exoneration is produced by a device whose function is precisely to preserve the moral self-image, that feeling loses probative value as an indication that exoneration has occurred. The relief ceases to be evidence of itself.

A more serious objection must be faced before proceeding, for it threatens the distinction at its root. It will be said that imputational relief is not a second state, but decisional relief accompanied by a false belief about imputation; that there would therefore be a single psychological phenomenon and a proposition believed about it. The objection would be decisive if the only proof of duality were the difference between feeling and believing.

The proof is another, and it has already been partly given: the double dissociation shown above demonstrates that each of the two reliefs occurs without the other, which no belief about a single state could explain. The difference of the defeaters makes the point exact. Two states are distinct when they are dissolved by distinct operations. Decisional relief disappears when the agent is obliged to decide: withdraw the directive, give him back the alternative, and the weight of deliberation returns in full, even if he remains convinced that the responsibility would lie with someone else. Imputational relief disappears through an entirely different operation: it suffices to show him that the transfer did not occur, and the weight returns without any part of the deliberation having been given back to him — the decision remains taken by another, and he answers for it. Each survives the suppression of the other, and neither is a condition of the other. They are not, therefore, a state and a belief about it, but two states with their own conditions of satisfaction.

From this follows the consequence that separates the two planes of analysis. Decisional relief is a primarily psychological and organizational phenomenon: it describes what the agent was spared from doing, and its occurrence is verifiable without anything being decided about imputation. Imputational relief is constituted by a normative claim, for whoever experiences it implicitly asserts that he is not the instance from which the act receives its moral qualification. This claim can be true or false, and it is assessable. It is because it contains an assessable claim that it belongs to moral philosophy, and not only to psychology — and that is why it remains to be asked, further on, whether it holds.

THE RESIDUE: ANSWERING FOR THE RECOGNITION

There remains the question that imputational relief left open and that the previous section made legitimate: does the claim contained in it hold? Does the subjective relief correspond to a real reduction of moral responsibility?

Smith (2015) provides the most direct route for addressing it. On her conception, to be responsible for something is to be subject to the demand to offer reasons that express or explain one's own evaluative judgement; responsibility is *answerability*, the property of being someone to whom the question why can legitimately be addressed, and from whom an answer can be expected that reveals the judgement on the basis of which one acted. What makes an agent responsible is not having chosen in an unconditioned way, but the fact that his conduct expresses evaluations that are his own and for which he can be called to account — which is why the conception dispenses with any requirement that the agent be the ultimate origin of himself.

Applied to the case, this conception produces a precise result. Asked why he acted, the Inquisitor's subject answers: because I was ordered to. The answer is true and, in many contexts, sufficient. But it does not close the series, for a second question immediately follows, which bears not on the act but on the answer: why did you treat that order as a reason? And, more specifically, why did you treat that order, with that scope, issuing from that source, as a sufficient reason for this?

The decisive point is that this second question cannot be answered by delegating it. To recognize an authority is itself an evaluative judgement, for it implies assessing that the source is competent, that its scope covers the matter, and that its directives merit pre-emption. This judgement is the agent's, it expresses his evaluations, and it remains subject to answerability. Delegating the first-order judgement does not delegate the second-order judgement that authorized the delegation, because delegating the latter would require a third judgement, and so on indefinitely. There is a structurally non-transferable residue, and it lies not in the decision, but in the recognition.

Arendt reached a similar conclusion by an entirely different route, and the record matters because hers is the best-known analysis of the relation between obedience and responsibility. Reflecting on the controversy generated by her report on the Eichmann trial (Arendt, 1963), she maintained that the question addressed to those who carried out orders should never be why they obeyed, but why they supported, and she proposed eliminating the word obedience from the moral and political vocabulary, since it describes as passivity what is in fact adherence (Arendt, 2003).

The adult who obeys in political and moral matters consents, and the question about consent is the same one formulated here under the name of the judgement of recognition. The limit of this contribution must, however, be marked, for the case she analyses is the inverse of the one in this article: her Eichmann is banal, and banality is the absence of thought, whereas the Inquisitor is lucid, knows the secret, anticipates the objections, and assumes the curse. The

category of thoughtlessness does not reach him, and it is precisely the case not reached by it that is examined here.

Before proceeding, four terms must be fixed that the rest of the section will keep apart, for current discussion uses them interchangeably, and this interchangeability is the origin of a good part of the confusions dealt with here. Justification is a property of a reason or of a directive, and the corresponding question is whether there are valid reasons. Answerability is the agent's position before the question about the evaluative judgement that his conduct expresses. Blameworthiness is the fittingness of an unfavourable response to that judgement. Liability to consequences is the attribution, generally conventional, of the burden of bearing the outcome, and it is the only one of the four constituted by convention, from which it follows that it is the only one redistributable by convention: this is what vicarious civil liability, the liability of the principal for the acts of the agent, and insurance do.

The four vary independently — a directive may be justified and the agent blameworthy for having followed it for the wrong reason; an agent may owe an answer without being blameworthy; liability to consequences may fall upon someone who has nothing to answer for —, and the separation proposed here converges with Shoemaker's (2015) warning that "responsibility" names distinct dimensions — attributability, answerability, and accountability to sanction —, which must not be treated as a single quantity. The Inquisitor operates precisely on the improper passage between them.

As for the interlocutors: Smith provides answerability and the reason why citing the authority does not satisfy it, for the authority is precisely what reasons are being asked for; what she does not provide is the treatment of the case in which the capacity to formulate the second question has itself been produced, or suppressed, by whoever benefits from the answer. Fischer and Ravizza (1998) provide the requirement that the mechanism through which the agent acted be properly his own and remain responsive to reasons, which allows the next question to be formulated precisely: would the mechanism of obedience, as the Inquisitor constructs it, be responsive to reasons?

It is at this point that the current answer goes wrong, and the error deserves to be exposed because it is seductive. It is argued that an agent whose obedience was constructed to be insensitive to any counter-reason displays, by that very construction, diminished imputability — whence the success of the Inquisitor's project would produce less responsible agents. The inference is hasty, and the distinction that undoes it is Edwards's distinction between natural inability and moral inability (Edwards, 1957, I.4), announced in the introduction and now applied.

Natural inability is the absence of the faculty and results from an impediment external to the will; it excuses. Moral inability is the faculty present but held captive by a disposition that is the agent's own; it does not excuse, because it is not an impediment to guilt, but guilt itself in the form of captivity, and to excuse it would be to absolve sin by pleading sin.

Applied to the Inquisitor’s subject, the distinction forces the separation of two components that the psychological description keeps together. The first is volitional: he does not want to doubt, because doubt has been presented to him as a fault and because obedience brings him peace. The faculty of judging remains intact, and what immobilizes it is his own disposition. This is moral inability, and it does not mitigate. The second is epistemic: he lacks the secret, deliberately withheld, and therefore the material without which the judgement of recognition could not be correctly formed.

It would be tempting to treat the second component as a case of natural inability, but the assimilation does not withstand examination: natural inability is the absence of a faculty, whereas ignorance is the absence of an object on which the faculty may operate, and Edwards would classify as moral the ignorance that the agent himself cultivates. The appropriate instrument is another, and it can be formulated without confessional vocabulary: blameworthiness varies according to the epistemic conditions that the agent could have altered with the diligence that could be required of him. Where he could have altered them and did not, the lack of knowledge is itself an object of blame; where no diligence that could be required would have altered them, blame for what depended on them does not hold. It may be noted, without adopting them, that two traditions converge on this point by different paths: scholastic moral theology formulated it as the distinction between vincible and invincible ignorance, and the Reformed idiom prefers to speak of guilt proportional to the light received, resisting the first formulation because it seems to establish a class of inculpable agents. The convergence matters; the choice between the two does not, because the criterion employed is the analytical one stated above, and it grades blameworthiness without at any moment extinguishing the obligation.

The secret withheld by an institution that holds the monopoly of interpretation and qualifies inquiry as a fault constitutes an epistemic condition that the subject could not have altered, and to that exact extent — to that extent, and not to the full extent of the relief experienced — blame is mitigated. The refusal to raise the question, on the other hand, when the means were available and what was lacking was the disposition to use them, is a blameworthy omission. The two conditions coexist in the same subject, in proportions that vary with circumstances, which confirms gradability without conceding exoneration.

The result is asymmetrical, and the asymmetry is the point. The construction of obedience insensitive to counter-reasons does not, by itself, produce non-imputable agents; it produces morally incapable agents, who answer. It mitigates imputation only to the exact extent of the epistemic deprivation effectively imposed, and that extent is much smaller than the relief the agents feel. Between the relief experienced and the exoneration due there is a distance, and it is this distance that the Inquisitor’s project exploits.

Two consequences follow from this result; the first closes this section and the second will occupy the next. The first concerns the foundation of obligation, and it requires an argument of its own. A common way of sustaining it consists in identifying a determinate position as constitutive of the moral agent and deriving the duty from it; this route, however, obliges one to defend that

identification, and the available identifications depend on presuppositions that the declaration of neutrality made at the outset prevents us from invoking. Here, therefore, no position is identified as fundamental, and nothing is asserted about the ultimate origin of norms. Only the formal thesis is maintained that, whatever the position in question, the obligation derives from it and not from the present capacity to fulfil it — and this formal thesis admits an independent argument.

The argument is as follows. Suppose the contrary, namely, that obligation follows present capacity, in such a way that the loss of the latter extinguishes the former. Since incapacity can be produced by the obligated party himself, or accepted by him, it would follow that self-incapacitation would be a mode of exoneration, and every norm would contain, in its very formulation, the instruction for revoking it. A norm that cancelled itself by the expedient of becoming impossible to comply with would not be a norm, but a recommendation conditional on the convenience of its addressee.

The argument is not conclusive by itself, and the rejoinder it provokes — that responsibility can be attributed by tracing back to the earlier moment at which the agent had the capacity — will be faced in the section on objections, under the name of tracing, where the suspicion of circularity that the notion of position may raise is also answered. It is nonetheless worth anticipating in what sense tracing does not make the thesis dispensable, so that the thesis is not left exposed in the meantime: tracing supposes a capacity that existed at some moment, and it decides on blame, not on the scope of the norm — two restrictions that the section on objections develops. It suffices to retain the conclusion here: obligation is addressed to whoever occupies the position to which the norm is directed, and supervening incapacity alters what can be demanded of the obligated party without altering the fact that he is the obligated party.

The law recognizes this independently, and the convergence is significant precisely because it is practical and not doctrinal: whoever incapacitates himself through a culpable disposition answers for what he does in that state, whereas incapacity not assumed mitigates or exempts. This is what is expressed, in criminal law doctrine, by the doctrine of *actio libera in causa* and by the treatment given to voluntary intoxication, which does not exclude imputability. The legal rule reproduces, in another vocabulary, Edwards’s distinction, and both point to the same criterion — what excuses is not the absence of power, but the absence of power that the agent neither produced nor assumed.

Applied to the case, the result is immediate. The subject recognized the authority and, in recognizing it, placed himself in the position of one who deferred. Answerability for the recognition is addressed to that position, and not to the capacity he may retain after having deferred, on pain of successful deference becoming the means of its own immunization. Responsibility for the act is graded according to epistemic conditions, as has been shown; being subject to answerability is not graded by the same expedient.

The limiting case remains, and it must be faced. Consider the subject formed entirely within the system, who never performed an act of recognition because he never had at his disposal

a vocabulary in which legitimacy could be called into question. Of him it will not be said that he placed himself in a position, and answerability for the recognition does not reach him. The point is conceded: blameworthiness, in this case, is substantially reduced. What is conceded must, however, be delimited — the claim concerns imputation among human agents for a determinate act, which is the register in which the entire article moves, and nothing is asserted about ultimate responsibility, a question that exceeds it and on which the traditions diverge.

The concession does not favour the Inquisitor, and the reason lies in the structure of his own project. He does not promise to produce agents incapable of judging; he promises relief to agents who judge. The entire device analysed — the anticipatory justification, the available moral description, the narrative of duty — presupposes a conscience that demands reasons and that suffers for not having them, and a subject without that demand would need no relief at all. The project is therefore self-limiting: to the exact extent that it produced the subject formed entirely within the system, it would cease to produce what constitutes its product. The case in which the thesis of position does not apply is the case in which the Inquisitor no longer has anything to offer, and the subjects of the poem — who pray, fear, sin, and seek someone to absolve them — are not that case. This is the reason why the article recognizes degrees without concluding in favour of transferability. The limit of this rejoinder must, however, be noted: it proves something about the character's project, and not about the extent of the concession. The subject formed entirely within a system is not a literary artifice, and the total institution is a documented phenomenon; outside the poem, therefore, the concession reaches more cases than within it. Measuring that reach is a question of fact, and the article does not decide it.

THE NON-CONSERVATION OF RESPONSIBILITY

The second consequence announced above makes it possible to formulate the final objection to the Inquisitor in his own terms. The moral economy he proposes presupposes that responsibility is a conserved quantity: a determinate load which, once transferred to whoever accepts it, ceases to weigh on the one who handed it over. His entire offer rests on this presupposition, and it is false. Normative responsibility does not obey a law of conservation. That the Inquisitor accepts to answer for everything does not subtract what the subjects have to answer for; it merely adds what he has. And the subjects' responsibility diminishes exactly to the extent that their agency has been compromised by deprivation not assumed, and not to the extent that someone has offered to carry it. The offer is, in this precise sense, ineffective in what it promises and effective only in what it conceals: it does not reduce imputation, it reduces the perception of imputation.

It is worth making explicit why it fails, and the theory of imputation offers the instrument. Two ways are distinguished in which an imputation reaches someone who did not perform the act: by real solidarity, when the one who answers shares in the nature or the position of the one who acted; and by a merely external arrangement, when a pact attributes to someone the consequences of an act in which he does not participate. The second way, on its own, does not bind justly, which

is why explanations of imputation by representation are invariably led to resort to the first (Feinberg, 1968; Dietrich; Dietrich, 2026a). The distinction is structural and does not depend on the case that historically motivated it. It should be noted that it does not correspond to the vocabulary of Kutz (2000), which is that of participatory intention: what his analysis of complicity provides is the reason why someone answers for a collective harm in which he participated, and not the reason why someone would answer for a harm in which he did not participate — which confirms, by another route, that the Inquisitor's offer finds no support in that literature.

Applied here, the distinction shows what the Inquisitor's offer lacks: he proposes a transfer of the second type and claims the effects of the first, for nothing binds him to the subjects except his own willingness to assume the burden, which is a unilateral act. And the reach of this unilaterality is fixed by the distinction established above. A unilateral act can add liability to consequences to whoever performs it, and can even assume that of third parties where convention authorizes it; it subtracts from no one blameworthiness or answerability, which are not constituted by convention and are therefore not redistributed by it. This is precisely what the Inquisitor promises: he does not offer to pay the subjects' fine; he offers that they cease to be the ones whose task it is to answer. The offer therefore has the grammar of a substitutive imputation and not the structure that would make it intelligible. It is for this reason, and not because of any moral severity on the part of the analyst, that the relief it produces is real and the exoneration it promises is null.

THE LUCID FANATIC AND THE SILENCE

It remains to examine the position of the Inquisitor himself, and it is the most instructive, because in him imputational relief appears in the form that resists every internal objection. The Inquisitor is coherent. He knows the secret, recognizes that he works with the other spirit, accepts the curse, harbours no illusions about what he does, and nonetheless maintains that he does well. There is, in his discourse, no contradiction that can be pointed out, nor any self-deception that can be denounced: he satisfies all the internal conditions of reflective endorsement.

His is, exactly, the case of the reflectively impeccable and yet mistaken agent, whose existence shows that no merely immanent criterion distinguishes the one who is right from the one who merely basks in his own image, because immanent criteria measure internal properties whereas correctness is an external relation (Dietrich; Dietrich, 2026b).

The point is not ornamental. It explains why the episode cannot be resolved by internal criticism. To show the Inquisitor that he contradicts himself is impossible, because he does not contradict himself; to show him that he is mistaken about his motivations is gratuitous, because he declares them. If his position is to be assessed, the assessment will have to be external — and this is exactly what justificatory closure aims to prevent. The monologue constructed so as not to

admit any reply is the discursive form of immanence: a space of reasons into which nothing from outside can enter.

It is against this background that the silence becomes legible, and the warning about attribution made at the beginning should be repeated here, for it is in this scene that it matters most. The figure who falls silent and kisses is the Christ of Ivan's poem, a literary construction by a character who declares that he does not believe, and not a proposition about the nature of divine authority; what the scene warrants are conclusions about forms of authority, not about Christology. With this proviso: the Christ of the poem does not reply discursively; he approaches and kisses the old man on his bloodless lips (Dostoiévski, 2008, Book V, ch. 5). Two hasty readings must be avoided. The first treats the silence as defeat or as absence of argument, which ignores that, in the economy of the poem, it is the silent figure who holds authority. The second treats the kiss as a rhetorical victory, a gesture that would refute the discourse by non-discursive means — which would reintroduce through the back door the same logic of closure, only with the sign inverted.

The reading proposed here is another: the gesture does not close; it gives back. It does not supply the Inquisitor with the answer that would dispense him from answering, and it leaves him exactly where he had tried not to be, in the position of one who must answer for himself. The text records the result precisely, and the record is the argument: the kiss glows in his heart, and he remains faithful to his idea. The old man remains an agent, and the answer remains his. Nothing was decided in his place, not even against him. It should be noted that the reading does not require the gesture to succeed: the text is compatible with its failure, and that is probably how Ivan wants it. What it requires is only that the gesture leave the old man in the position of one who has to answer — and it is this, and not conversion, that the final record of the chapter attests.

The opposition can then be formulated in the terms constructed throughout the article. On one side, an authority that seeks to produce conformity through justificatory closure, occupying in its entirety the space in which the subject would find his reasons. On the other, an authority that addresses the subject without occupying the place of his answer and that, for that reason, preserves him as a participant in the practice of justification, and not as its addressee. It is not necessary to decide here theological questions about the nature of true Christian authority; the philosophical function of the scene is narrower and safer, and consists in showing that the opposite of domination is not the absence of authority, but a form of authority that preserves the interlocutor as an agent capable of answering. It is remarkable that this formulation converges, by entirely distinct paths, with the requirement of reciprocal justification in Forst and with the conception of responsibility as *answerability* in Smith: what the scene figures in gesture is what both state in concept, namely, that a moral agent is someone from whom an answer is expected, and that to take that position away from him is not to relieve him, it is to diminish him.

The final symmetry of the chapter confirms the reading, and it is Ivan himself who points it out. When the poem is over, Alyosha kisses his brother in silence, and Ivan protests, laughing,

that this is plagiarism. Here too the reply is not a counter-discourse that would close the discussion, but a gesture that gives Ivan back his own question.

OBJECTIONS AND DELIMITATION OF SCOPE

It is objected that the thesis of position is circular, since it defines the position through an act of the agent and thus reintroduces the capacity it sought to exclude. The objection confuses two capacities. It is not denied that positions are acquired through voluntary acts, nor that the acquisition requires capacity when it occurs; what is denied is that the obligation, once the position is established, continues to follow the capacity to fulfil it. Whoever enters into a contract must be capable of contracting, and the contractual obligation is not extinguished when the debtor becomes incapable of performing. It is the second dependence, and not the first, that the argument rejects.

It is objected, with greater force, that the argument against the capacity thesis proves less than it appears to, since no serious defender of that thesis dispenses with tracing: responsibility can be attributed by tracing back to the earlier moment at which the agent had the capacity and culpably destroyed it (Fischer; Ravizza, 1998). The device is conceded; it is legitimate and resolves the majority of cases. Two observations, however, show that it does not make the thesis of position dispensable. Tracing presupposes a moment at which the capacity existed, and says nothing about those who never had it. And it answers a different question: it determines whether the agent is blameworthy for the violation, not whether he is under the norm. An agent may be obligated without being blameworthy for failing to comply, and it is this subsisting obligation that explains why his situation is described as an excuse, and not as the absence of duty.

It is objected that obedience under coercion really does diminish responsibility, and this is correct, within the limit that the previous section fixed. Coercion, fear, imposed ignorance, and material dependence alter imputability when they amount to a deprivation not assumed by the agent. What the article denies is the automatic inference from the relief felt to the exoneration due, and it denies it with Edwards’s distinction: where the faculty remains and it is the disposition that immobilizes it, there is no excuse.

It is objected that the argument would presuppose libertarian freedom. It does not. The entire course of the argument was formulated in terms of responsiveness to reasons, ownership of the mechanism of action, the capacity to answer for evaluative judgements, and obligation deriving from position — notions that do not require indeterminism and that remain compatible with current compatibilist architectures.

It is further objected that the subject is to a large extent the product of whoever addresses him, and that holding someone responsible for dispositions that another formed in him is unjust. The objection is serious and is not answered by denial. It is answered by the distinction already employed and by the gradation already conceded. The formation of the disposition does not remove the faculty of judging, and it is the faculty, and not its origin, that moral inability supposes

to be present; what the heteronomous origin alters is the measure of blame, and that measure was expressly tied, above, to the epistemic conditions that the agent could not have altered with the diligence that could be required of him. The objection, therefore, does not establish exoneration: it establishes that mitigation may be greater than ordinary intuition admits, which this article concedes, and concedes without touching answerability, which is not graded by the same expedient.

It is objected that the distinction between the two reliefs is merely psychological, and that moral philosophy would have nothing to say about it. The objection was faced above through the argument from defeaters, and it is worth only recording its conclusion: imputational relief is not a feeling to which a belief is added, but a state whose very constitution involves a claim about who answers for the act — and claims of this kind are assessable, which suffices to place it within the domain of moral philosophy.

With the concessions gathered together, it is worth saying what remains standing, for their sum might suggest a thesis in retreat. It was conceded that the pre-emption of part of deliberation is legitimate, that tracing resolves the majority of cases of self-inflicted incapacity, that liability to consequences is transferred by convention, that the character's anthropology may be true, that the agent never formed as a judge has reduced blameworthiness, and that the heteronomous origin of dispositions may widen the mitigation due. None of these concessions touches the result. It remains that imputational relief is a state distinct from decisional relief, with its own defeaters; that it contains an assessable normative claim; that this claim is false where the incapacity is moral; that justificatory closure takes hold in knowability and acceptance, and not in the applicability of reasons; and that a third party's willingness to answer adds what he has to answer for without subtracting what the agent has. The concessions delimit the scope of the thesis; they do not replace it.

Finally, the scope of the conclusion must be delimited. It does not follow from this article that deference to authority is in general illegitimate, nor that the structure described here typically characterizes this or that institutional tradition; claims of that kind would depend on empirical evidence that the text neither produces nor invokes. Nor does anything here settle the question of the extent to which each concrete agent was deprived of the conditions for judging, for that is a question of fact, and a variable one. What is established is more modest and more precise: that imputational relief contains a normative claim distinct from decisional relief, that this claim is assessable, that it is false when the incapacity at stake is moral and true only to the extent of the epistemic deprivation imposed, and that a third party's willingness to answer for the act changes none of this, because responsibility is not a conserved quantity.

FINAL CONSIDERATIONS

The aim has been to show that obedience produces two effects that the literature treats as a single block, and that separating them alters the assessment of both. Decisional relief spares the

agent the work of deliberating and asserts nothing about imputation. Imputational relief offers him a moral description under which he ceases to understand himself as the normative source of the act, and this description contains a claim that may be true or false. That the two are distinct states, and not one state and a belief about it, is established by the difference of their defeaters: each survives the suppression of the other.

Once the claim is assessed, the result is asymmetrical. It is false where the faculty of judging remains and what immobilizes it is the agent's own disposition, for this is moral inability and does not excuse; and it is true only to the extent of the epistemic conditions that the agent could not have altered with the diligence that could be required of him. Between the relief experienced and the exoneration due there is, therefore, a distance, and it is this distance that the offer of transfer exploits.

It remains to determine what each interlocutor provided and what each did not, for it is from this reckoning that the position maintained here emerges. Bandura provides the mechanisms by which relief is produced; he does not provide, nor does he claim to provide, any conclusion about imputation, and from the efficacy of a mechanism no normative assessment whatsoever follows. Arendt provides the replacement of the question why one obeyed by the question why one supported; she does not provide the treatment of the lucid agent, for her diagnosis is the absence of thought. Smith provides the demand to answer for evaluative judgements and the reason why citing the authority does not satisfy it; she does not provide the treatment of the agent whose capacity to judge has been configured in advance. Shoemaker provides the warning that responsibility names distinct dimensions; he does not provide the hierarchy among them when they diverge. Fischer and Ravizza provide the requirement of a mechanism that is one's own and responsive to reasons, and the device of tracing; they do not provide the distinction between being obligated and being blameworthy, which tracing presupposes without stating. Edwards provides the distinction between natural and moral inability, by conceptual argument and not by doctrinal authority; he does not provide the formal thesis about position, which this article had to establish by an argument of its own.

It remains to indicate where this points. The offer examined is not peculiar to the religious setting of the case chosen, and it reappears whenever an instance supplies simultaneously the directive and the moral description under which it is to be understood. The final formulation may be stated thus: the Inquisitor's project is not completed when individuals simply obey, but when they find in the authority itself the justification through which they can understand obedience as duty and control as care. The most effective form of domination does not silence conscience; it supplies it in advance with the reasons with which it will be able to absolve, explain, or dignify its own conduct. That these reasons relieve it is a fact about conscience; that they exonerate it is a claim about the world, and it is false.

REFERENCES

ARENDT, Hannah. *Eichmann in Jerusalem*: a report on the banality of evil. New York: Viking Press, 1963.

ARENDT, Hannah. Personal responsibility under dictatorship. In: ARENDT, Hannah. *Responsibility and judgment*. Ed. Jerome Kohn. New York: Schocken Books, 2003.

BAKHTIN, Mikhail. *Problemas da poética de Dostoiévski*. 5th rev. ed. Trans., notes and preface by Paulo Bezerra. Rio de Janeiro: Forense Universitária, 2010.

BANDURA, Albert. Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, v. 3, n. 3, p. 193–209, 1999.

DIETRICH, Pablo; DIETRICH, Maurília Gabriela da Silva. A responsabilidade como arquitetura: participação, imago Dei e os limites do fundamento único. *Revista Eletrônica Amplamente*, Natal/RN, v. 5, n. 3, p. 938–946, Jul./Sept. 2026a.

DIETRICH, Pablo; DIETRICH, Maurília Gabriela da Silva. O espelho e o critério: Feuerbach como constitutivista do valor e os limites da crítica antropológica da religião. *Revista Eletrônica Amplamente*, Natal/RN, v. 5, n. 3, p. 947–960, Jul./Sept. 2026b.

DOSTOIÉVSKI, Fiódor. *Os irmãos Karamázov*. Trans. Paulo Bezerra. São Paulo: Editora 34, 2008.

EDWARDS, Jonathan. *Freedom of the will*. Ed. Paul Ramsey. New Haven: Yale University Press, 1957. [Originally published 1754.]

FEINBERG, Joel. Collective responsibility. *The Journal of Philosophy*, v. 65, n. 21, p. 674–688, 1968.

FISCHER, John Martin; RAVIZZA, Mark. *Responsibility and control*: a theory of moral responsibility. Cambridge: Cambridge University Press, 1998.

FORST, Rainer. Noumenal power. *The Journal of Political Philosophy*, v. 23, n. 2, p. 111–127, 2015.

FROMM, Erich. *Escape from freedom*. New York: Farrar and Rinehart, 1941.

KUTZ, Christopher. *Complicity*: ethics and law for a collective age. Cambridge: Cambridge University Press, 2000.

RAZ, Joseph. *The morality of freedom*. Oxford: Oxford University Press, 1986.

RAZ, Joseph. The problem of authority: revisiting the service conception. *Minnesota Law Review*, v. 90, p. 1003–1044, 2006.

SANDOZ, Ellis. Philosophical anthropology and Dostoevsky's "Legend of the Grand Inquisitor". *The Review of Politics*, v. 26, n. 3, p. 353–377, 1964.

SHOEMAKER, David. *Responsibility from the margins*. Oxford: Oxford University Press, 2015.

SMITH, Angela M. Responsibility as answerability. *Inquiry*, v. 58, n. 2, p. 99–126, 2015.

TSANG, Jo-Ann. Moral rationalization and the integration of situational factors and psychological processes in immoral behavior. *Review of General Psychology*, v. 6, n. 1, p. 25–50, 2002.