Is a Coherent Theory of Religious Freedom Possible?

Steven D. Smith
Is a theory of religious freedom possible? It's obvious that we can and do talk about and argue about issues of religious freedom—school prayer, aid to parochial schools, and so forth. But a "theory" entails something more than "talk," or ad hoc argumentation. ¹ More specifically, we would treat talk as falling short of being a "theory of religious freedom" on either of two grounds that are relevant here.

First, we sometimes distinguish between a "theory" and something else that we might describe as a compromise or "modus vivendi." A modus vivendi doesn't give us an internally consistent set of principles capable of generating answers to questions of religious freedom; it is more in the nature of a negotiated, and perhaps messy, truce.

Second, a position does not qualify as a theory of religious freedom if it begins by preferring one (or some subset) of the competing religious and secular positions, and then proceeds to spell out what that preferred position does and doesn't allow. Suppose I contend, for example, that the law should permit teacher-led school prayer but not compulsory baptism; and when asked to explain these conclusions I argue that they follow from the best interpretation of Catholic theology, or perhaps Mormon or Muslim theology, and that this particular theology is the truest or best one available. I think we do not consider this sort of position to be a theory of religious freedom. On the contrary, this sort of position is as a historical matter entirely familiar, and it is just what religious freedom is supposed to save us from.

¹ Thus, I don't think arguments that certain arrangements regarding religion are better or worse than other arrangements necessarily implies that a theory of religious freedom does or must exist, as some scholars suggest. See Christopher L. Eisgruber and Lawrence G. Sager, Unthinking Religious Freedom, 74 Tex. L. Rev. 577, 591-92 (1996).
I. WHY THERE CAN BE NO THEORY OF RELIGIOUS FREEDOM

My view is that all argumentation about religious freedom will be disqualified from being a "theory of religious freedom" on one or the other of these grounds.¹ It may be helpful if I give the reason for my view in summary form, and then elaborate.

I think the establishment and free exercise controversies that we are familiar with present one aspect of a more universal problem, which we might call the problem of "the spiritual and the temporal."³ "For the flesh lusteth against the Spirit, and the Spirit against the flesh: and these are contrary the one to the other . . . ."⁴ The enduring problem is to determine how these matters stand in relation to each other. Does the spiritual take priority over the temporal, or vice versa? Is the spiritual more real, or more authoritative, than the temporal—earth being merely a footstool for heaven? Or, conversely, is the spiritual merely derivative, or epiphenomenal, or perhaps merely a delusion?

Problems of religious freedom present one manifestation of this conflict within the realm of politics and law. These problems may involve competing claims to authority advanced by spiritual and temporal institutions. Or they may involve spiritual claims made by individuals that conflict with more general temporal interests. And when such spiritual-temporal conflicts arise, possible responses or resolutions can be understood as falling into three general categories, which I will describe as "spiritual primacy," "temporal primacy," and "dualism." The first two categories offer the possibility (at least in the abstract) of "theory," or of principled resolutions of conflicts—but not of resolutions that can usefully be described as respecting "religious freedom." The last category—dualism—is capable of recognizing a place for religious freedom, but it does not offer the possibility of principled resolutions of conflicts.

¹. For a lengthier and somewhat differently structured argument for this conclusion, and one that considers "originalist" as well as "nonoriginalist" arguments, see Steven D. Smith, Foreordained Failure: The Quest for a Constitutional Principle of Religious Freedom (Oxford U. Press, 1995).
³. Different people would use different vocabularies, of course; they might distinguish between the soul with its goods and the body with its appetites, or between nature and grace, or between our welfare in this life and in the next.
⁴. Galatians 5:17.
A. THE PRIMACY OF THE SPIRITUAL

I want to try to clarify these general observations by considering a concrete example of each of these responses. So let me start with the category of “spiritual primacy.” In this view, the temporal is subordinate to, or perhaps a subset of, the spiritual. And it seems to follow that spiritual-temporal conflicts should be resolved by applying spiritual criteria.

As an example of this approach, consider the medieval papacy’s conception of government and law, as interpreted by Walter Ullmann.5 Ullmann explains that

the papacy, in common with medieval doctrine and literature, held that the individual’s activities cannot be separated into more or less well defined categories.... Christianity seized the whole of man—man was whole and indivisible: every one of his actions was thought to have been accessible to the judgement by Christian norms and standards.6

Those Christian norms and standards were ultimately directed toward a spiritual and indeed otherworldly end: the salvation of the soul. But it did not follow that the affairs of this life were unimportant. On the contrary, “while the end of this [Christian] society and of its members was in the other world, the terrestrial life was nonetheless of fundamental importance in achieving this other-worldly aim, that is, salvation. The principle of indivisibility embraced the life in this as well as in the other world....”7

Within this spiritual conception of life, the Church was responsible for administering the Christian norms that governed all earthly activities.8 Consequently, the papacy regarded the

---

5. I should say that although Ullmann’s depiction, like all historical interpretations, is debatable; I’m using it here not so much to make historical claims as to illustrate one possible approach to our problem.
7. Id. at 34.
8. Precisely because this Church was an entity that existed on this earth its direction concerned therefore the doings of its members on this earth. The vital point was that these earthly activities of the Christians must be directed by Christian norms, which meant that they must be guided, orientated, directed.

Id. at 34.
Id. at 73.
Id. at 35.
Church and the pope as possessing jurisdiction to direct all the affairs of Christendom. Within this comprehensive jurisdiction, Ullmann argues, all other authorities (including both bishops and princes) were subordinate to papal authority. "Power, that is, jurisdiction, was concentrated in the pope, who handed part of it on to the bishops, part of it to kings and emperors, and so forth." In sum, "the secular prince ... [was] a necessary, auxiliary organ ... instituted by divinity to assist the pope in his government."

This conception provided an intellectual framework within which conflicts between religious and secular authorities could be adjudicated. To use our terms, that framework allowed for a "theory." I think, because it contained inclusive principles that were accepted by the competing interests as those interests were understood at the time. Of course, kings and emperors often resisted the popes' claims, and in the realm of power and politics their resistance was often successful. But in the realm of theory the secular rulers were severely handicapped because they themselves embraced the inclusive premises on which the papal claims rested.

Thus, Ullmann explains that "it would be wholly erroneous to think that these principles were, so to speak, imposed upon kings and princes."

No king or emperor ever objected to the papal theme that his kingdom was entrusted to him by God: on the contrary, it was the kings themselves who, quite independent of, and uninfluenced by, the papacy had adopted this standpoint.

Of course, a king could claim that he received his power from God directly, rather through the intermediary of the pope. The emperor Henry IV made just this claim in his famous dispute with Pope Gregory VII. But in an officially Christian world this assertion seemed weak. The New Testament re-

---

9. Id. at 55. In this allocation of powers, the specific function of the secular prince was the suppression of evil, including heresy, by force. Id. at 64-66, 79-82. It was for the Church to judge what was evil—full authority or sovereignty lay only with the pope, id. at 67, 72, 87—and for the prince to act upon and enforce that judgment.
10. Id. at 65.
11. It meant that each office-holder should fulfil the functions contained in his office, and no more. The king should not interfere in the functions of the bishop, because he was not created for this purpose; the archdeacon should not meddle with matters pertaining to the sheriff, and so forth.
12. Id. at 88.
13. Id. at 61.
corded Christ's conferral of power on Peter—and, by inference, on his successors—but what evidence was there of any independent divine conferral of authority on the king or the emperor? And what special competence could a worldly and perhaps illiterate prince claim in matters of scripture and Christian doctrine? Ullmann stresses that within the Christocentric worldview, papal claims to sovereignty were perfectly logical—indeed, virtually irrefutable. We might put the point more generally: If the primacy of a spiritual position is accepted within a community, then it is natural that the office or institution responsible for preserving and interpreting that spiritual position should enjoy ultimate authority within the community.

Consequently, the possibility of a secular authority independent of the Church awaited the emergence of a dualistic worldview in which the "temporal" was freed from its subordination to the "spiritual." Opponents of papal authority tried to develop this position. Ullmann explains: "What the dualists aimed at in their opposition to the papacy was the ascription of autonomous and indigenous character to the 'temporal'. . . . In this way it was believed that the monarchy of the king could be saved: in temporal matters the king was to be the monarch, in spiritual matters the pope."14

This dualist talk begins to sound familiar to modern ears accustomed to the theme of "separation of church and state." But in the medieval climate of opinion the dualist view made little headway for the understandable reason that the spiritual and the temporal did not seem severable. Dualism, Ullmann explains, "contradicted not only the Pauline doctrine, which the papacy had made its own, but also the principle of totality or indivisibility, which in itself was the message of Christianity, seizing as it did the whole of man and the whole of his activities without splitting them up into different compartments."15

In sum, in a world in which competing parties concurred in accepting a set of premises that treated the temporal as subordinate to the spiritual, it was possible to develop a theory regulating the relations between, to use our terms, church and state. And this theory could contemplate—indeed, insist upon—a division of functions between "secular" officials such as kings and religious officials such as bishops. Still, it would seem a little

14. Id. at 97. "[D]ualism of government was to be the panacea of royal govern­ments from Henry IV in the Investiture Contest, who actually coined the term and in­vented the idea, down to the Reformers and beyond." Id.
15. Id.
strange to describe this theory as a theory of religious freedom. The label seems inappropriate not because the theory allowed secular authorities no freedom—in fact the popes never tried to dictate every decision made by kings and emperors—but because a system under which secular authorities are auxiliaries to the Church and have just as much power as the Church assigns to them is not the sort of arrangement we have in mind when we talk about "religious freedom." On the contrary, we would call such a system a "theocracy." And we are accustomed to treating "theocracy" not as a version of, but rather as the antithesis of, "religious freedom."

B. THE PRIMACY OF THE TEMPORAL

I've been describing the medieval papacy's conception of government, as depicted by Ullmann, as an example of a system that gives priority to the spiritual. What might be an example of the opposite position—that is, of a system that treats the spiritual as subordinate to or a subdivision of the temporal? You might suspect that the answer is very close at hand—that as legal scholars we actually inhabit just such a system. Maybe so, but to avoid controversial characterizations, let me adopt the law teacher's trick of describing a partly hypothetical culture that will illustrate my point—one that you can recognize or not as you choose.

Imagine then a community in which most people ultimately believe in and care only for the temporal. They believe, perhaps, that this life is all we have: "When we die, we die." This community also believes that there is no guiding intelligence or overall purpose or design in the cosmos. So the purpose of people—and hence of governments established of, by, and for the people—is to promote the welfare of human beings in this life. The community and most of its members are committed solely to the pursuit of temporal values and interests. Of course, this community might still care about "spiritual" things in a watered down sense of the word—it might use the adjective to describe things like opera and poetry, for example—but it has rejected religious entities or notions such as God, the soul, grace, the supernatural, and life after death.

But lest we solve the problem of religious freedom too quickly simply by eliminating religion, we can suppose that a residual commitment to the spiritual lingers on in several ways. First, a few members of the community retain a genuine faith in
God, the soul, life after death, and the divine authorship of scripture. These religious believers are widely regarded as peculiar and backward, and at least among the more educated, their faith is viewed as “irrational superstitious nonsense.” Even so, they are capable of causing unpleasantness. Moreover, in view of the community’s egalitarian pretensions these religionists are thought to be entitled in some ill-defined sense to “equal concern and respect.”

In addition, some members of the community who do not actually believe in God or the soul still feel a degree of selective affection toward religious practices and traditions. They may think some traditions are useful in the upbringing of children. They expect that as adults these erstwhile children will come to understand the purely pragmatic function of the traditions—while of course continuing to pass on the traditions to their own children. Others may find religious practices soothing (Gregorian chants, maybe), or aesthetically attractive, or useful on special occasions for expressing emotions like hope or grief. There are also agnostics of an antiquarian bent who find that religious traditions and rituals help them to preserve ties to the past, and perhaps to maintain a sense of personal or communal identity.

In short, “religion” continues to exist in various senses even in this devoutly temporal community. But would there be any room, or any reason, to give special honor or legal status to religious freedom? Of course, the community might recognize and in a sense respect practices that for some purposes are classified under the heading of “religion.” The community might protect these practices, that is, because they implicate temporal interests. If government interferes with religious belief or exercise, the interference might injure people’s peace of mind or sense of identity. Frustrated religionists might become uncivil or even violent. For reasons like these, the community might conceivably carve out an area of human activity called “religion” and afford it different or special legal treatment. And it’s even conceivable, if unlikely, that this special treatment could be successfully codified in terms of some principle or theory. People who advocated such a position might claim that they had articulated a theory of religious freedom.

Still, there are good reasons to question both this possibility and this characterization. Notice first that although the hypothetical position could be called a “theory of religious freedom,” it does not offer any protection to religion as religion, or because of its character as religion. Rather, the theory protects religion as temporal human activity and because that activity is thought to affect temporal interests. To put the point differently, from the community’s standpoint, it is in a sense merely fortuitous that the activities and beliefs which affect the temporal interests in question happen to be religious activities and beliefs. So we might doubt that it is helpful, or accurate, to attach the label of “theory of religious freedom” to an account that does not even count the religious character of a belief or activity as relevant in itself or for its own sake.

This question about truth-in-labeling points to a related practical objection: The temporal interests invoked by the theory will probably not correlate cleanly with “religious” activities and beliefs, and so it will seem both illogical and imprudent to make the theory’s application coextensive with what is for other purposes called “religion.” Suppose, for example, that a theory is based on the value of self-realization. It holds, perhaps, that government should not interfere in religious choices because these choices are too closely linked to a person’s sense of who she is, or to her very identity; and self-definition or self-realization are temporal interests that the community should respect. The point about labeling suggests that it would be more accurate to call this a “theory of self-realization,” not a “theory of religious freedom.” The related practical point is that self-realization probably does not correlate precisely, or even very closely, with religion. For many people, choices that are not religious may be central to self-realization, while choices that involve religion may not always affect self-realization in any essential way. So the logic of the theory suggests that it should abandon its claim to being a theory of religious freedom not only in name but also in practical application.

18. The draft exemption cases illustrate this progression. They started with statutory language that plainly exempted only persons whose objection to war arose from a set of beliefs centered on faith in a “Supreme Being”—i.e., God. But in the modern climate of legal opinion this sort of focus seemed incongruous or unacceptable. Hence, the cases first expanded the definition of what would count as a religious objection and then made it clear that the expanded definition would apply even to a belief that the conscientious objector himself did not regard as a religious belief. Justice Harlan went even further, explicitly asserting that the availability of the exemption could not depend on...
One response to this objection might argue that although "religion" is not identical to temporal interests like personal autonomy, still for practical purposes religion is a good proxy for those interests. But this argument seems dubious at best. Particularly in a pluralistic culture where "religion" takes a variety of radically diverse forms, it seems *prima facie* implausible that the amorphous category of "religion" would correlate even approximately with any particular temporal interest or set of interests.

In this situation, how would we account for continuing efforts to theorize about religious freedom? If a given community already has a longstanding commitment to religious freedom derived from earlier events and other grounds (perhaps by now largely forgotten), then it would be tempting to interpret temporal arguments for religious freedom as *post hoc* and less than persuasive rationalizations for a commitment that the community has inherited but no longer fully grasps. And a prescription that might naturally follow from this interpretation is that the community should clean up its theory and practice in light of its current beliefs and values. One way to do this would be to dissolve religious freedom into other, more current theories and commitments—commitments to free speech, perhaps, or to equality.

In this way the community conceivably might bring its practice into line with what is for it a plausible theory or set of principles. And under other headings, like "free speech" or "equality," the community might continue to give legal protection to some beliefs and practices that the believers and practitioners themselves regard as "religion." But the purified temporal community would not now claim to have, or to operate according to, any "theory of religious freedom."

---


20. Leading constitutional scholars who take this approach include Lawrence Sager and Christopher Eisgruber, see, e.g., Sager and Eisgruber, 74 Tex. L. Rev. 577 (cited in note 1), and Ira C. Lupu, Keeping the Faith: Religion, Equality, and Speech in the U.S. Constitution, 18 Conn. L. Rev. 739 (1986).

C. DUALISM

The positions I’ve talked about thus far have been monistic in the sense that they have regarded either the spiritual or the temporal as primary and encompassing, with the subordinate term being viewed as a subdivision of the primary or preferred category. I have argued that at least in the abstract both perspectives offer the possibility of “theory” prescribing principled resolutions for some of the disputes that we treat under the headings of “religious freedom,” or “church and state”; but neither perspective actually recognizes the value of “religious freedom” in a meaningful sense.

That value is associated, rather, with a dualist position—one that regards both the spiritual and the temporal as independently valuable. This sort of perspective, with both its promises and problems, is reflected in John Locke’s “A Letter Concerning Toleration.” Locke acknowledges the importance both of spiritual interests—the salvation of the soul—and of temporal or “civil interests,” or “things belonging to this life,” such as “life, liberty, health, and indolency of body; and the possession of outward things, such as money, lands, houses, furniture, and the like.”

And spiritual and temporal values are independent of each other; property, for example, is a good not because it will contribute to our eternal salvation, but because it is valuable in the here and now.

Based on this dualistic view, Locke argues for a division of responsibility between the church and the state. The church’s function is to care for the salvation of souls, and in such matters of salvation the state has no legitimate concern. The state’s function, rather, is to protect and promote the civil interests. In this way, Locke concludes that church and commonwealth are “perfectly distinct, and infinitely different from each other” and that “[t]he boundaries on both sides are fixed and immovable.”

Of course, dualistic thinking and imagery have pervaded discussions of religious freedom in this country. The notion surfaces again and again: Church is church and state is state; they are separated by a high and impregnable wall, and both are happier for the split. The dualist view is attractive because it seems

23. Id. at 35.
to recognize the value of religious freedom in a more meaningful sense than do either of the more monistic views.

In addition, presentations like Locke’s even seem at first to allow for a theory of religious freedom. The task of theory is to determine which interests belong to the spiritual domain and which belong to the temporal, and in this way to draw a line of demarcation between the spheres. Once that line has been fixed, then when questions of religious freedom arise we need only examine the interests closely to see which side of the line they fall on.

Upon closer examination, though, the pleasant prospects offered by the dualist view dissolve. One way to consider the problem is to ask from whose perspective it can be said that the church’s concern is solely for the salvation of souls, while the state’s responsibility is for this-worldly, “civil” interests. If this is the way the church itself and the state itself understand their respective domains, then a nice harmony of interests seems possible. Indeed, it becomes difficult to explain why religious freedom and church-state relations ever created such a ruckus in the first place. On the other hand, if the church or the state do not understand their roles in this way, so that Locke’s description of roles is merely his own (or perhaps the state’s, or the church’s) view of what the division of responsibilities ought to be, then no principled resolution has been achieved. Instead, the pretense of such a resolution is in reality merely the imposition of terms by a dominant party employing a deceitful description of what the other party cares about.

In fact, churches historically have not understood their concern to be limited to the salvation of souls in the next life. Later in his essay, Locke acknowledges that both religion and the state are vitally concerned with issues of morality. The concession effectively negates his earlier claim that the spiritual and temporal spheres are “perfectly distinct and infinitely different from each other”—a claim which is essential to the possibility of a dualistic or separation principle.

A different way of highlighting the false promise of dualism is to ask whether the different interests of the spiritual and the

24. For example, many churches have regarded the achievement of education, social justice, and civil rights as major religious objectives. Even more importantly, religions rarely separate the concerns of this life from those of the next; they have typically taught that the way a person lives here will powerfully affect the disposition of the soul after death.

temporal are wholly different, or whether their respective spheres of interest are different but overlapping, as in a Venn diagram. If the spheres of interest were wholly different, then a principled resolution of conflicts should be easily attainable; indeed, the harder problem once again would be to explain why such conflicts ever arise in the first place. But if the spheres overlap (as they almost certainly do), then it is natural to suppose that conflicts, and hence questions of religious freedom, will arise within the area of overlap. And within that area, the dualist promise of a principled resolution cannot be fulfilled. Since both the spiritual and the temporal make a claim, any solution will necessarily reject one of those claims by giving priority to the competing perspective.

And precisely because it recognizes that the spiritual and the temporal are both valuable, and that they are independently valuable (as opposed to one being derived from or a subset of the other), dualism cannot dictate which perspective should prevail. If the spiritual and the temporal are both real and autonomous, in other words, then there is no more encompassing principle to which they are both subordinate.

Any actual controversy involving religious freedom can serve to illustrate this controversy. Suppose that Native Americans assert a religious duty to use peyote in religious rituals, while a state insists on banning the use of harmful drugs, which in the state's view include peyote. The use of peyote in this context affects both a spiritual interest and a temporal interest. A dualist view would observe that both interests are independently valuable; neither is simply derivative of or reducible into the other. And the response to this dualist observation is, "Yes. That's exactly why we have a conflict." In this context the spiritual and temporal domains overlap, and in order to resolve the conflict, someone will have to treat either the spiritual or the temporal as primary and the other interest as subordinate.

This objection does not imply that the dualist view is wrong, I think, or that the dualist view cannot value religious freedom. The point is merely that dualism itself cannot provide any theory or principle defining the proper scope of religious freedom.26 On the contrary, the occurrence of actual controversies demonstrates that a point has been reached where the contribution of

26. Does the absence of a "principle" mean that judicial review in this area would necessarily be "unprincipled" and therefore illegitimate? For a tentative discussion, see Steven D. Smith, Unprincipled Religious Freedom, 7 J. Contemp. Legal Issues 497 (1996).
dualism is insufficient to produce peace. At that point a decision-maker will be forced to give priority to either the spiritual or the temporal perspective, and we will once again be faced with the same problems that afflict the monistic positions.

II. THE INNUMERABLE CITIES

The moral of this story is that our concept of religious freedom gets its meaning within a dualist framework. But a dualist framework cannot honestly offer any useful principle, or support any theory, for adjudicating the problems we consider under the heading of religious freedom. So to the extent that religious freedom flourishes within a dualist system, it will necessarily reflect an accommodation or a modus vivendi among spiritual or temporal interests, not a deduction from theory or principle.

This conclusion should not be surprising. On a personal level, most of us are quite accustomed to negotiating between the spiritual and the temporal. True saints (if there are any) or true atheists (if there are any) may manage to achieve monistic peace. Meanwhile, most of us try to serve both masters, and so we live in tension, making choices that are at times arbitrary, intuitive, ad hoc, not regulated by any encompassing principle or theory. We might say that we “balance” the spiritual and the temporal, except that the metaphor of balancing suggests more precision than we really experience. It might be more accurate to say that we “juggle” the spiritual and the temporal, acknowledging that many of us are clumsy jugglers who are often forced to stoop and pick up the pieces. As we’ve been told, the “double-minded man is unstable in all his ways.”

On a more global level, the moral of the discussion is also a familiar one—famously presented in, for example, Augustine’s City of God. Religious believers are citizens of two cities, Augustine explained—an earthly city and a heavenly one. Both cities seek “peace.” This apparently common goal might make it seem that the cities are nicely harmonious; and indeed a limited harmony may be achievable. But in fact by “peace” the two cities understand quite different things. “The earthly city, which does not live by faith, seeks an earthly peace, and the end it proposes... is the combination of men’s wills to attain the

27. James 1:8.
28. “[A]s this life is common to both cities, so there is a harmony between them in regard to what belongs to it.” Saint Augustine, City of God 19:17 (Marcus Dods, trans. & ed., Hafner Publishing Co., 1948).
things which are helpful to this life." The heavenly city takes advantage of this peace, but it also understands that "the perfectly ordered and harmonious enjoyment of God and of one another in God" is what "alone can be truly called and esteemed the peace of the reasonable creatures..." These different conceptions and different ultimate attachments and aspirations sometimes come into conflict, and when that happens "the heavenly city has been compelled... to dissent, and to become obnoxious to those who think differently."

In the final analysis, therefore, the heavenly city "lives like a captive and a stranger in the earthly city." The heavenly city can seek benign terms of captivity, or perhaps a sort of truce, from the earthly sovereign; it cannot hope for genuine understanding and union.

Augustine's analysis remains valid for our modern society, I think, except that our more developed pluralism—in the realms both of religion and of government—may make the description of two cities seem a little quaint. For us, it seems, there are not two but rather many cities—many conflicting faiths, worldviews, lifestyles, cultural and political allegiances. This riotous pluralism merely underscores the conclusion suggested above: We may aspire to achieve a modus vivendi among these competing positions, but it is a mistake to suppose that there can be any encompassing theory or "neutral principle" to regulate their interaction.

29. Id. (emphasis added).
30. Id.
31. Id.
32. Id.