

CAN CONVENTIONS SUPPORT THE LEGAL INTERPRETATION OF SCRIPTURE?

Lara Buchak

Department of Philosophy, Princeton University, USA
Correspondence email address: lbuchak@princeton.edu

ABSTRACT: Scripture seems to command actions that our modern moral sensibilities find immoral, which poses a problem for those who take Scripture to be the word of God. In response, Amir Saemi has proposed the Legal Interpretation of Scripture, according to which Scriptural commands are legal rather than moral. On this view, Scripture does not always prescribe the morally best action, but instead prescribes the rule that yields the morally best attainable outcome, given feasibility constraints. Using a formal account of coordination, I evaluate this solution, and show that resolving the problem requires Scriptural legislation to function not through reward and punishment but instead by creating mutual expectations about what others will do. I further argue that we should not understand moral progress as increased virtue; rather, moral learning explains how legislating non-optimal rules can lead agents to converge on the optimal ones. The resulting account refines Saemi's proposal while vindicating its central idea. I close by noting some remaining worries.

KEYWORDS: Problem of divinely prescribed evil; revelation; scriptural interpretation; coordination and convention; moral learning

Introduction

In his recent book, *Morality and Revelation in Islamic Thought and Beyond*, Amir Saemi tackles an important problem that he calls the “new problem of evil”: the problem that Islamic Scripture seems to command actions that our modern moral sensibilities find immoral. For example, Islamic Scripture seems to claim that men are in charge of women, that women's testimony is less trustworthy than men's, and that men are allowed to beat their wives in certain situations. Other examples include allowing slavery and killing apostates. While this book is written in the Islamic context, it is

clear that many religions face a version of this problem: their Scriptures, on the seemingly most plausible interpretation, command actions that seem to modern readers to be immoral.

Saemi notes two standard solutions to this problem. One solution is linguistic reinterpretation: when we pay more attention to the historical and linguistic context of Scriptural commands, we discover that the actual commands are different than they seem. But matching up Scriptural commands with modern sensibilities requires a radical departure from their apparent and historically understood meaning. Another solution is contextualization: while the commanded actions are wrong in our current context, they were not wrong in the radically different context in which they were issued. For example, in seventh-century Arabia – when the Qur’an was revealed – it was common to be excessively violent towards women; and so the command to only beat women in particular circumstances was a *restriction* on permissible violence towards women. The problem with this solution, according to Saemi, is that even if restricted wife-beating is better than unrestricted wife-beating, it is still not *morally justified*, even in the context in which the command was issued.

We are left with two options. One is to hold that Scripture takes priority over our moral judgments. Saemi argues that this view (the Scripture-first view) is the dominant view in the history of Islamic philosophy. The other option is to hold that our moral judgments take priority over Scripture (the ethics-first view). This is Saemi’s view, and he spends much of the book arguing for it.

On the ethics-first view – under the assumption that Scripture is divinely inspired – we have to explain why God would command or allow immoral actions. To answer this challenge, Saemi proposes a view of Scripture that he calls the Legal Interpretation, which says, roughly speaking, that Scriptural commands are legal commands but not moral commands. In Saemi’s words:

Actions prescribed or permitted by the best interpretation of a Scriptural passage are *legally* obligatory or permitted, according to religious law. But God or the Prophet may legislate legal obligations or permissions that deviate from *moral* ones. (Saemi, p. 208)

To explain why God or the Prophet might legislate immoral actions, Saemi draws on Scott Shapiro’s “planning” conception of the law, according to which the law is a *social plan* designed to solve moral problems (Saemi, p. 217). A key fact about planning is that it is irrational to make plans that are infeasible – that we know cannot be carried out. Therefore, a legislator is constrained by which options are feasible, and must sometimes legislate rules that are morally non-optimal, if there are no morally optimal rules that are feasible.

One example of this comes from the ethics of war (Saemi, p. 220). It might be true that, in a war, if one side is just and the other is unjust, then it is morally acceptable to kill both combatants and “active” civilians on the unjust side, but it is not morally acceptable to kill anyone on the just side. However, it would be infeasible to enforce these principles: in practice, most countries will think they are on the just side, so the rule “kill only those on the unjust side” would lead to a lot of deaths. The Geneva

Convention holds that combatants can kill each other, but not civilians. Even if this is not the morally optimal rule, it leads to less death than alternative rules.

Saemi argues, then, that the rules of Scripture are not morally optimal, but that it is nonetheless good for God or the Prophet to legislate them. Not only are they the best rules that are feasible at the time of legislation, they can help move society towards the morally optimal rules. In the particular case of wife-beating: according to Saemi, Scripture allowed it in certain circumstances, because although it is morally impermissible, allowing it in those circumstances was morally better than the other feasible options (e.g. allowing it in all circumstances), and this legislation can help Islamic society gradually move towards a complete prohibition on wife-beating.

In this article, I will examine these claims more closely. I will first identify four conditions that must be met in order for Saemi's account to explain why Scripture prescribes actions that are morally non-optimal. To see whether these conditions are in fact met, I will introduce a formal account of coordination. I will then show what function legislation must have in order for it to successfully legislate the non-optimal rules but not be able to legislate the optimal rules. Roughly: it must not function by introducing rewards and punishments, but instead by creating expectations about what others will do and about what others will expect others to do. Given the ethics-first solution, the account must also say why the optimal solution is available in the modern day, even without Scripture legislating it. The best way forward for this solution, I argue, is not to hold that people have gotten more prone to virtue; instead, when we examine the process of moral learning, we come to see why legislating a non-optimal solution can sometimes lead people to eventually converge on the optimal one. I take my solution to be a friendly amendment to and expansion of Saemi's proposal. I close with some potential worries.

Four Conditions

To explain his account of religious laws, Saemi draws on a story by Ibn Tufayl (1105-1185) (Saemi pp. 210-214). In this story, a person named Hayy knows all the moral truths. He travels to a nearby island, full of vicious people who follow the teachings of a prophet sent by God, but these teachings surprisingly don't align with the moral truths. Hayy tells the people of the island the moral truths, but because they are vicious, it actually makes them behave worse. Hayy thus discovers that non-ideal laws can sometimes have better outcomes than ideal laws, in particular when people are vicious. In Saemi's words:

Religious laws, in their capacity as legal laws... are the laws that serve in nonideal circumstances to make communal life possible... When most people act immorally and find correct moral principles repugnant, one can't solve social problems of the society by legislating the correct moral principles. The lesson we learned from the tale of Hayy is that there could be reasons to maintain the permissibility of immoral laws. Sometimes the best way to solve the social problems of a corrupt

community is through legislating morally incorrect principles. (Saemi, pp. 213, 221)

According to Saemi, nascent Islamic society is like the island in the story. Legislating morally correct principles would have made people behave worse than legislating morally non-optimal principles; and legislating morally non-optimal principles made people behave better than they would in the absence of any laws.

This account has something in common with the contextualization solution: it points out that Scriptural commands, at the time they were issued, were morally better than existing behavior. But unlike that solution, it doesn't hold that the contents of the commands are moral truths. According to the Legal Interpretation solution, restricted wife-beating is *not*, in fact, morally permissible – but issuing the command to do it is nonetheless the best God can do.

To make this account more precise, assume there are three possible norms a society could adopt. The first is a deeply inegalitarian norm, which allows men to beat their wives even for trivial reasons. The second is a moderately inegalitarian norm, which allows men to beat their wives only in extreme circumstances. The third is the egalitarian norm which never allows men to beat their wives. And assume that the egalitarian norm is in fact the morally best norm, followed by the moderately inegalitarian norm, followed by the deeply inegalitarian norm.

Saemi's account rests on several key claims. The first two:

PAST VICE: In the absence of Scripture, men in nascent Islamic society would adopt the deeply inegalitarian norm.

LEGISLATION EFFECTIVENESS: If Scripture legislates the moderately inegalitarian norm, then men in nascent Islamic society will adopt the moderately inegalitarian norm.

Together, these claims imply that by legislating the moderate norm, Scripture made things better than the status quo. The third claim:

LEGISLATION INEFFECTIVENESS: If Scripture legislates the egalitarian norm, then men in nascent Islamic society will fail to adopt the egalitarian norm, and will instead continue to adhere to the deeply inegalitarian norm.

Along with the first two claims, this claim implies that legislating the moderate norm in fact leads to the best societal outcome, even if it is morally sub-optimal.

Saemi also seems to hold that the moderate norm is not the final resting place: that present-day people could instead adhere to the egalitarian norm. First, Saemi holds that the laws don't just improve the status quo – they also move us towards the morally optimal solution (Saemi, p. 221). Second, he holds that we currently understand that sexism is wrong, and so permission to beat one's wife even in extreme circumstances would be a bad social plan in our society today (Saemi, p. 222). Finally, he holds that we should be guided by the ethics-first view in our present society, rather than the

Scripture-first view: presumably, following our own ethical ideals will result in a better outcome than following Scripture.

Therefore, Saemi's final claim is that:

PRESENT VIRTUE: In the absence of Scripture, men in present-day Islamic society will adopt the egalitarian norm.

The key question is whether these four conditions can be simultaneously vindicated. To answer this question, we need to look more closely at individual motivation and rationality – and, in particular, at how individual preferences interact to produce the relevant facts about what people will do, both in the absence of Scripture and in the presence of various Scriptural commands. In particular, we must examine the general nature of coordination and convention, and how norms are maintained and changed.

Coordination and Convention

In his influential book *Convention*, David Lewis uses simple game theory models to analyze what conventions are, how they solve coordination problems, and how they are formed and maintained. A coordination problem is a situation in which two or more individuals are trying to choose actions in light of interests that coincide. For example, we are each trying to decide which side of the road to drive on, and our primary aim is to drive on the same side of the road as each other, to avoid accidents. We can describe our situation using a game matrix, where the numbers represent the “utility” payoffs to each of us for each combination of actions, Row's payoff followed by Column's (Figure 1). The magnitude of the numbers is inessential: what matters is their relationship to each other.

	You Drive on the Left	You Drive on the Right
I Drive on the Left	10, 10	0, 0
I Drive on the Right	0, 0	10, 10

Figure 1: Choice of Where to Drive

In this scenario, we don't care which outcome we coordinate on, as long as we coordinate. In other scenarios, we might have a preference about which outcome we coordinate on, even if coordination is more important to us than otherwise satisfying our preferences. For example, consider the choice of what to wear to work. We are most interested in wearing the same type of clothing, so as to not be embarrassed; but beyond that, we both find sweatpants more comfortable than khakis, and we both find khakis more comfortable than suits. Our situation can be represented in Figure 2.

	You Wear Sweatpants	You Wear Khakis	You Wear a Suit
I Wear Sweatpants	30, 30	15, 10	0, 0
I Wear Khakis	10, 15	20, 20	5, 5
I Wear a Suit	0, 0	5, 5	10, 10

Figure 2: Choice of What to Wear to Work

The numbers represent three important facts. First, holding fixed what the other person does, your highest payoff can be found by matching their action. Second, if you come close to matching them, but don't match them exactly, you still get a moderate payoff: if you wear khakis, then I am only mildly embarrassed by wearing sweatpants or a suit, since each of these "isn't too far off" from khakis. Third, being far away from matching the other person is the worst of any of the options: if you wear a suit when they wear sweatpants, or vice versa, you will be very embarrassed.

What each of us does depends not just on our own payoffs, but on our expectations of what the other will do. If I expect you will wear khakis, for example, then that is a reason for me to wear khakis. But it is also important that you expect that I expect you will wear khakis. If you expect that I expect you will wear a suit, then you will expect that I will wear a suit, and so you will wear a suit. In this way, coordinating on the same action both requires and is supported by higher-level expectations.

We say that we have "solved" a coordination problem when we end up in one of the "coordination equilibria": one of the outcomes in which neither of us could do better by either one of us unilaterally acting otherwise (both wearing sweatpants, khakis, or suits).

One way to solve coordination problems is explicit agreement: if we say to each other that we will both wear khakis, then we will both expect that the other will wear khakis, we will both expect that the other will expect us to wear khakis, and so forth. But coordination problems can be solved without explicit agreement. They can also be solved by precedence: if we both wore khakis to the office last week, we will have higher-order expectations that we will wear khakis. They can also be solved by salience: if we have no other reasons to coordinate around any particular equilibria, we might each choose the one that "sticks out." An important way for an equilibrium to stick out is for it to be the unique equilibrium that everyone prefers.

Coordination problems don't just take place between two people; they also occur in populations. Lewis says that a *convention* arises among members of a population in recurrent situations of the relevant form when everyone conforms to some regularity and expects everyone else to do so, and when these facts are common knowledge.

Importantly, once a convention arises, it is *rational* for each individual to stick to it, and *irrational* to deviate. For if I can reasonably expect that others will do their part, then I do best by doing my part too. Thus, conventions are maintained because it is individually rational for each person to do their part. It is also possible for conventions to change, similarly to how they are formed in the first place. If I come to expect that others will do their part in a *different* convention, then it will be rational for me to do my part in this new convention, rather than the old one. The rational force of

conventions thus comes from each individual's preferences, plus each individual's higher-order expectations about each other's behavior.

According to the Legal Interpretation, Scriptural commands help people coordinate around an outcome that is superior to what they would do if left to their own devices. We can clearly see this in the example of commands telling people when to pray, if it is important for everyone to pray at the same time (Saemi, pp. 229-230). But this example has two specific features: we are neutral about what times to coordinate around (as we are about which side of the road to drive on); and when to pray is morally neutral. Can the theory of convention also explain why Scripture would legislate morally non-optimal actions rather than morally optimal ones? More generally, can we give a model of the situation that satisfies all four of the conditions outlined in the previous section? I turn to this question now.

A Model of the Problem

Let us assume that each individual man has three options: inegalitarian actions, moderate actions, or egalitarian actions. Now, recall the four claims that Saemi's solution rests on. Given PAST VICE, let us assume that – in the absence of any facts about what others will do – men in nascent Islamic society preferred inegalitarian actions to moderate ones to egalitarian ones. But let us also assume that men care more about coordinating – taking the same actions as each other – than about satisfying their personal preferences. This is exactly analogous, in terms of each man's preferences, to the situation of dressing for work (distasteful as the comparison may be!). As in the workwear example, we will assume that if men don't do the same actions as their fellow men, they fare better insofar as they are closer to doing what their fellows do rather than farther away. The assumed situation is represented in Figure 3.

	You are Inegalitarian	You are Moderate	You are Egalitarian
I am Inegalitarian	30, 30	15, 10	1, 0
I am Moderate	10, 15	20, 20	5, 5
I am Egalitarian	0, 1	5, 5	10, 10

Figure 3: Choice of Actions in a Marriage¹

These numbers represent men's preferences, and the example of Hayy and the island certainly implies that these preferences arise from viciousness. But I note that this is not essential to the model itself: preferences could, for example, be "moral" preferences about what would be best for the family unit. For example, the preferences of someone who thinks inegalitarian actions are morally better and wants to act in accordance with moral good, but personally dislikes carrying out the inegalitarian actions, could be represented by these numbers.

¹ This represents a two-player game, but it is easy to extrapolate to the multi-player version.

It is natural to think that in the absence of Scripture, everyone will take the inegalitarian actions, and the inegalitarian norm will thus become a convention. For given that everyone prefers the inegalitarian actions, in the absence of reason to coordinate around anything else, everyone will form higher-level expectations that everyone will take the inegalitarian actions.

Finally, let us assume that the correct *moral* ordering is Egalitarian $\succ$ Moderate $\succ$ Inegalitarian. And let us imagine we are in the position of *legislators*, and we are trying to get people in this society to adopt the *morally best* convention we can get them to adopt.

Vindicating LEGISLATION EFFECTIVENESS and LEGISLATION INEFFECTIVENESS

There are two ways in which legislation can produce conventions. I argue that the first way cannot vindicate both of the conditions about effectiveness, but the second way can.

A Failed Model: Legislation Changes Payoffs

One possible function of legislation is to *change the payoffs* by introducing new consequences. In the case of state laws, I might be thrown in jail if I don't comply. In the case of workplace edicts, I might be fired if I don't comply, or given a good performance evaluation if I do. In the case of laws given by God, I might be punished in the afterlife for disobeying, or rewarded in the afterlife for obeying. This is naturally modeled as drastically increasing the utility of the legislated act (or, equivalently, drastically decreasing the utility of the non-legislated acts). For example, if moderate egalitarianism is legislated, then men's payoffs increase by 100 if they choose *moderate*:

	You are Inegalitarian	You are Moderate	You are Egalitarian
I am Inegalitarian	30, 30	15, 110	1, 0
I am Moderate	110 , 15	120 , 120	15 , 5
I am Egalitarian	0, 1	5, 15	10, 10

Figure 4: Effects of Legislating Moderate, if Legislation Changes Payoffs

If this is how the law functions, then LEGISLATION EFFECTIVENESS is obviously satisfied: whatever is legislated becomes the preferred action, no matter what others choose, and so will be chosen by each man. Unfortunately, this view of how legislation works faces serious problems in meeting the remaining two conditions.

Let us begin with LEGISLATION INEFFECTIVENESS. Notice that if Scripture has the relevant effect on preferences *and* the egalitarian actions are available, then Scripture could simply legislate egalitarianism:²

	You are Inegalitarian	You are Moderate	You are Egalitarian
I am Inegalitarian	30, 30	15, 10	1, 100
I am Moderate	10, 15	20, 20	5, 105
I am Egalitarian	100, 1	105, 5	110, 110

Figure 5: Effects of Legislating Egalitarian, if Legislation Changes Payoffs

Thus, an account that relies on the law changing payoffs needs to explain why the egalitarian norm is *unavailable*.

There are two ways in which an action can be unavailable. The first is that individuals are physically incapable of carrying it out. To use our above example, it might be that clothing companies are simply not making suits; thus, “wear a suit” isn’t an option for me in this sense. This can’t describe the situation here, however: even if some ingredients for an egalitarian marriage may have been physically unavailable in nascent Islamic society, always refraining from beating one’s wife was not *physically* unavailable.

The other way in which an action can be unavailable is that it can be difficult to carry out given your limited vantage point. This is true, for example, of the advice to “buy low and sell high” when investing in the stock market. Even if this action is physically available to you, you can’t know ahead of time which stocks will go up and which will go down. This is also true of the example from the ethics of war: “kill only those on the unjust side” is advice that many people cannot follow, since everyone will assume the other side is unjust. Importantly, though, “never beat your wife” isn’t difficult in this sense either: it isn’t epistemically difficult to follow this rule.

The problem with the idea that legislation changes payoffs is that it explains the power of the law to bring about the moderate norm, but it doesn’t explain its *powerlessness* to bring about the egalitarian norm, and Saemi’s solution requires both.

This view of legislation also struggles to vindicate PRESENT VIRTUE, at least in the absence of further explanation. If men continue to think that following the law will result in rewards or punishments, then they will continue to choose the moderate acts. If they no longer think that – if they revert to their original preferences – then there are only two possibilities. If they think that others think that following the law will result in rewards or punishments, or if they think that others will think that others will think this, etc., then they will believe that the moderate convention still holds, and they will choose moderate acts. If they *don’t* think this, then they will revert back to the inegalitarian norm. Either way, there isn’t any reason to think that they will adopt the egalitarian norm.

² There is another possibility, according to which following the law has only mild consequences – enough to move people to choose Moderate but not Egalitarian. This solution seems ad hoc; in any case, it still struggles to vindicate PRESENT VIRTUE, as we will shortly see.

A Successful Model: Legislation Produces Expectations

We saw that if the function of the law is to introduce rewards and punishments, then it cannot vindicate some of the conditions necessary for Saemi's solution. But this isn't the only way that legislation solves coordination problems. Another important function of the law is to change people's expectations, including higher-order expectations about each other's expectations, by making the prescribed outcome salient. Rewards and punishment are neither necessary nor sufficient for this function of the law. When a law with this function successfully prescribes an act, it makes everyone believe that everyone else will perform that act; and makes everyone believe that everyone else will believe that everyone else will perform that act; and so forth.

This point allows us to see another reason that the law might be ineffective in getting people to coordinate around the egalitarian norm. It might be that most men won't believe that other men will take the egalitarian actions, even if Scripture legislates them. Why not? Maybe these actions seem too outlandish. Here are some mundane examples to illustrate. If a politician were to successfully pass legislation that we should all drive in the middle of the road, doing so seems so obviously inefficient that we might not comply. If we all get an email from the boss saying "tomorrow's dress code is to wear a Halloween costume," it might sound so wild that we each think it's a prank. Even if we were each inclined to entertain these courses of action, we will additionally wonder whether each other will: I'll think that you probably won't drive in the middle of the road, and I'll think that you will probably think the email is a prank. So we will fail to form stable higher-order expectations about doing the legislated actions.

Similarly, it might be that in nascent Islamic society, egalitarianism would have seemed like an obviously inefficient way to run a household, like driving in the middle of the road. Or it might seem like it couldn't really come from God, like the email couldn't have come from the boss, and so maybe men wouldn't have even accepted that it was Scripture. Even if each man would think it might be efficient or might be Scripture, each man would also fail to form the expectation that each other man would think this, and so forth. As a result, each man would fail to form stable higher-order expectations about doing the egalitarian acts - and would instead revert to the inequalitarian ones.

Thus, we have arrived at an account on which both LEGISLATIVE EFFECTIVENESS and LEGISLATIVE INEFFECTIVENESS hold: in nascent Islamic society, while legislation can effectively cause men to form stable higher-order expectations about everyone following the moderate norm, it cannot cause them to form stable higher-order expectations about everyone following the egalitarian norm.

Vindicating PRESENT VIRTUE

We have seen how three of the conditions on which Saemi's account rests can be vindicated. The final piece of the puzzle is to vindicate PRESENT VIRTUE. We need to explain how individuals in the present day will choose egalitarian actions *merely* as a result of thinking they don't need to follow Scripture in our current time – rather than as a result of, say, new revelation legislating a new convention.

The Contours of a Solution

One way in which the possibilities can stop seeming outlandish – so that higher-order expectations about doing them become possible – is through precedence of coordinating around “closer” actions. For example, if a workplace has several events in which unusual-but-not-crazy dress is legislated (e.g. everyone wear pink, everyone wear a funny hat), then each of the workers might be more apt to believe they could all coordinate around wearing Halloween costumes. Precedents that are closer on the spectrum to an “outlandish” convention can make people more apt to believe that other people *might believe* that something is a convention. In our example, following the moderate norm might make people less likely to dismiss the egalitarian norm as absurd, or as appearing to one's fellows as absurd.

Unfortunately, this explanation can't get us all the way to coordinating around the egalitarian actions. Closer precedents secure the *possibility* of higher-order expectations for the egalitarian actions, but we also need some way for these expectations to get off the ground – for men to believe that other men will *actually* coordinate around these actions. In short: we need the egalitarian norm to have salience or precedence or agreement. But since no one prefers it, and it is not otherwise salient or precedented or agreed to, then men will have no reason to form higher-order expectations about its being chosen. Therefore, *in the absence of a new authoritative command*, we cannot yet vindicate PRESENT VIRTUE.

One possibility, of course, is to hold that there is a new authoritative command, via new Scripture or the authority of tradition. But for Islam in particular, Scripture is supposed to be the final revelation.³

In order to vindicate PRESENT VIRTUE, then, we need to attribute *different preferences* to men in the current day than in our model.⁴ Of course, it's reasonable to think that men's preferences have in fact changed. But this fact seems puzzling if it relies on the assumption that men in nascent Islamic society were selfish and corrupt but men now are not. We shouldn't think that people in the present in particular are

³ In 33:40, the Qur'an describes Muhammad as “the seal of the prophets,” which is typically understood to mean that Muhammad is the final prophet. Of course, one can debate the meaning of this verse or the force of other authorities in issuing new commands.

⁴ We've only been considering the coordination problem among *men*. One other possibility is to model this as a coordination problem *among men and women*, and to think that women's preferences or bargaining power have changed. I think there is truth in this suggestion, but I do not have space to explore it here. Even if we incorporate these facts, egalitarianism is much more stable if men have changed their preferences too.

less mean, selfish, etc., than people in the past – to think this seems to be a kind of chauvinism about the present. And even if we do think that people, on the whole, have become more virtuous over time, an ideal explanation of moral development would say how Scripture’s non-ideal commands could *bring about* this development.

I now propose a view that explains how preferences could change over time, even if past men and current men are similarly prone to virtue and have the same actions available to them. Importantly, it will explain how preferences can change *as a result* of non-ideal legislation.

Learning through Exploration

We have been assuming that men know the value they attach to various actions. But it might be that they aren’t sure of the value of actions they haven’t tried or norms they haven’t followed. After all, how can you be sure you are correct about the value of the egalitarian actions if you haven’t tried them? We can incorporate this thought into our model.

When you’re certain about the payoff of one action and don’t have a lot of evidence about the payoff of another action, you face a tradeoff in deciding what to do: if you do the uncertain-payoff action, you might get a lower payoff, but you will also get more information about that action. Figuring out how to navigate this tradeoff – which is sometimes called the tradeoff between “exploiting” and “exploring” – is the purpose of what are termed “multi-armed bandit” models.⁵ Here’s an example. Assume you are choosing between several slot machines, each of which has a true distribution of payoffs. When you choose one, you get the payoff it in fact produces, and you also update your assessment of its payoff distribution, given the information that it produced that payoff. Ideally, you should spend most of your time playing the slot machine whose expected payoff you estimate to be highest, but you should also spend some time trying out the other machines, in case one of them in fact has a higher expected payoff. As a general rule, the less costly it is to try out an alternative – the closer your estimate of its expected payoff is to your estimate of the one you currently estimate highest – and the less evidence you have about it, the more often you should try it.

Let’s apply this idea to our problem. Because multi-armed bandit models generally assume there is a distribution of payoffs rather than a constant payoff, we will make that assumption in our model, so that the numbers in the figures represent *expected* utilities rather than fixed utilities. (Taking the inegalitarian actions might have a range of results, depending on the circumstances, or one could be unsure how to assess the results of the actions.) The key additional assumption is that the numbers represent men’s *current estimates* of the expected value of various actions, but that men might not know their true expected values. For example, when thinking about the egalitarian convention, men estimate that following it has a payoff of 10. But it could be that its payoff is much lower – it could be that egalitarianism will lead to family dissolution –

⁵ See, e.g., Sutton & Barto (2014: Ch.2).

or much higher – it could be that it will lead to greater love and harmony within the family.

We can again remain neutral about what these payoffs represent: they could represent individual men’s flourishing or enjoyment, or they could represent their moral assessments. Insofar as we are more drawn to the view that men (then and now) are corrupt, it makes more sense to hold that they represent the former; insofar as we are more drawn to the view that men (then and now) are trying to behave morally but don’t always have the correct view about what constitutes morality, it makes more sense to hold that they represent the latter.

Assume that the *true* payoffs of the conventions match the actual moral values (either because the payoffs represent moral assessments or because the moral actions are in fact the ones that lead to flourishing), so the diagonals look something like this:

	You are Inegalitarian	You are Moderate	You are Egalitarian
I am Inegalitarian	30, 30		
I am Moderate		40, 40	
I am Egalitarian			50, 50

Figure 6: True Values of Actions in a Marriage

We can fill in the values of non-diagonal outcomes however we like, as long as the individual preference component is higher the more egalitarian an action is; this is compatible either with thinking that conforming is still more valued than taking the preferred actions, or with thinking that the values are the same no matter what others do.

At every time, each man faces a choice about whether to pick the action that he currently thinks is best, or to try out an alternative in order to gather more information about it. When the inegalitarian convention is in effect, it is more costly to try out alternative actions, given one’s original estimates of their values. If your fellows are all following the inegalitarian norm, then you will probably do much worse by not following it, and it is unlikely that you will do better, for two reasons. First, because of the “personal” component of the alternative actions (how valuable they are in the absence of a convention): the risk of marital dissolution isn’t worth the possibility of greater love and harmony, given that the inegalitarian actions already look pretty good. Second, because of the “convention-following” component: even if the personal component is better than estimated, one will be severely punished for not following the convention.

However, once the moderate convention is in effect, three things happen. First, because the moderate norm is the convention, men perform the moderate actions a lot more often – thus getting a lot more evidence about these actions’ true value – and they therefore update their assessment of moderate actions in the direction of their true value. Second, men might hypothesize that the true value of the moderate and egalitarian actions are connected – if the former leads to more love and harmony, it is more likely the latter will too – so they will also update their assessment of egalitarian actions in the direction of their true value (though maybe not as strongly). Third,

because trying out the egalitarian actions when the moderate convention is in effect is not as costly as doing so when the inegalitarian convention is in effect, men will be more incentivized to try it out occasionally, thus, again, updating in the direction of its true value.

Thus, men will slowly learn the true distribution of value of each action pair, and perhaps also learn that each other learns this. As a result, *dropping* the moderate convention will result in men *adopting* the egalitarian convention, since they all prefer it (and know that each other prefers it etc.). But, importantly, this process wouldn't have gotten off the ground without adopting the moderate convention in the first place: in the absence of that convention, trying out alternatives to inegalitarianism is too costly.

Let me summarize the main idea. When the inegalitarian norm is in effect, there is very little incentive to deviate from it, even though other actions might turn out better, and so men are in the dark about the values of the other actions. When the moderate norm is in effect, men slowly learn that moderate actions are better than they might have thought, and also that egalitarian actions might be better than they might have thought. They also have incentive to experiment with egalitarian actions, and such experimentation eventually results in coming to learn that egalitarian actions are preferred. Once enough people have the correct assessment, the egalitarian norm becomes salient, in the absence of legislation to the contrary.

A key feature of the account I've outlined is that it explains how non-egalitarian laws *help our moral development towards egalitarianism*: our desires change as a result of what moderate norms make possible, namely, experimenting with egalitarian norms. Legal commands – including non-ideal ones – are essential to the process of learning ethical truths.

The resulting account satisfies all four conditions. It explains why past men could have followed the moderate convention but wouldn't have been able to follow the egalitarian convention; and it explains why *present* men are apt to follow the egalitarian convention, even in the absence of Scriptural guidance. Importantly, according to this solution, PAST VICE and PRESENT VIRTUE are true, but not because present men are in fact more prone to virtue than past men. Instead, PAST VICE is true because men in nascent Islamic society *lacked knowledge of what it's like to follow very different rules*; and, given the convention they were stuck in, they had no motivation to try out the actions that would produce that knowledge. On this account, most people (past and present) are virtuous in a specific sense: they are able to recognize good actions *when they do them*. Although the account I've given is somewhat at odds with the example of Hayy and the island, I think that proponents of the Legal Interpretation and the ethics-first solution should nonetheless welcome it: after all, the idea that people should defer to their own ethical commitments makes more sense if most people are somewhat virtuous than if a large group of people are vicious.

Remaining Worries

I've provided an account of how to vindicate the four claims on which Saemi's Legal Interpretation is based. I briefly summarize its elements. First, men in nascent Islamic society preferred to be less egalitarian rather than more, but cared more about following the same norms as each other than satisfying their preferences. Second, they were able to form higher-order expectations about following moderate norms, but not about following egalitarian norms, perhaps because these norms seemed inefficient or they believed others would believe they were inefficient. Third, following moderate norms allows men test out egalitarian norms, and they eventually come to realize they prefer them. Therefore, in the present day, men will follow egalitarian norms in the absence of Scriptural injunctions.

A few worries remain about the account sketched here, and about the Legal Interpretation as a solution to the new problem of evil. First, why didn't God create a world in which men have a more efficient way to learn which actions are better? After all, knowing the true commands sooner would result in less harm to women. The account provided holds fixed what men are like and considers which commands are optimal in light of their nature, but we could instead ask why they have this nature; this is an instance of the (old) problem of moral evil.

Second, as mentioned, Islamic Scripture is supposed to be the final revelation. There is something unsatisfying about the final revelation containing non-optimal commands, especially if we are now in a position to coordinate around the optimal commands.⁶ Indeed, if Saemi's story is correct, then it would certainly be less confusing if God were to issue new commands (or provide an expiration date for the old ones), especially given that some people think Scripture does describe the morally correct norms.

Perhaps a proponent of the Legal Interpretation solution could argue that the morally correct norms are only feasible *given that* the sub-optimal ones are recommended. The idea might be that each individual's moral journey must necessarily follow the historical journey described: a person early in their moral development is inclined to the bad actions and can only be moved by Scripture to the better but sub-optimal ones. By accepting Scriptural constraints, each individual will experiment with the morally correct actions and come to see that they are preferred – a fact that (a proponent of this view might claim) can only be learned through experience rather than Scriptural testimony. That is: that Scripture is a kind of scaffolding in our individual lives as well as in society. This is a fairly revisionary view of Scripture, but it is an available view.

Third, as we saw, if religious laws are to have the function they need to have to vindicate the solution, then their motivational force cannot be the force of reward and punishment; instead, it is the force of higher-order expectations. This also seems like a fairly radical departure from how Scripture is traditionally conceived. In addition, it makes it difficult to receive guidance from Scripture in the present day if all it tells us is the optimal way to resolve social problems in light of the feasibility constraints

⁶ See Farbod Akhlaghi's contribution to this issue.

produced by past societies. So this kind of view would need to say more about the role of Scripture in guiding religious people today.

Conclusion

The Legal Interpretation is a novel and interesting view of Scripture. But the devil (or God) is in the details. A solution to the new problem of evil must explain how Scripture can be efficient but not too efficient, and how past men could want the worst but present men could want the best. I've provided an account on which all these conditions can be satisfied. It rests on the recalcitrance of expectations when "outlandish" actions are recommended, and an account of moral learning where these actions are tried out when costs are not too high, and found to be the best. Whether this story is a fully satisfying religious account of Scripture is a further question, but I hope I've at least articulated an account that rests on plausible facts about humans and their development, and illuminated how non-ideal norms can play a role in discovering the ideal ones.

References

- Lewis, D. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Saemi, A. (2024). *Morality and Revelation in Islamic Thought and Beyond*. New York, NY: Oxford University Press.
- Sutton, R. S. & Barto, A. G. (2014). *Reinforcement Learning: An Introduction* Cambridge, MA: MIT Press.

How To Cite This Article

Buchak, Lara. (2025). "Can Conventions Support the Legal Interpretation of Scripture?" *AGATHEOS: European Journal for Philosophy of Religion*, Vol. 2, No. 3, pp. 52-67.

Copyright Information

This is an open-access article published under the terms of the Creative Commons Attribution License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.