Suppose that I have become convinced that being a moral actor is a goal independently worthy of pursuit. I may have spent a significant amount of time studying a range of ethical theories and have become well-acquainted with the recommendations of various flavors of utilitarianism, Kantianism, other deontological systems, modern theories of rationality, and possibly a handful of religious doctrines. Perhaps I am comfortable enough with the rules of many of these doctrines that in any action scenario, I feel confident in my abilities to answer questions of the form “what would an act utilitarian do?” or “what would a Kantian do?”. Yet when these recommendation systems come into conflict, I am at a loss for how to decide which action is the most good. Each of these systems forms a unique set of instructions for how to be a “good” person, yet they quite frequently do not align in their prescriptions. I might be able to delve deep into the literature of arguments comparing and contrasting the merits of these systems, but at the end of the day still believe that there are valid reasons to support and reject any system, and remain at a loss, failed to be convinced that one particular doctrine is the end-all-be-all for moral action. For any moral system I choose, I might encounter scenarios
where the prescribed action would result in a feeling in my gut that what I am doing is wrong. I may be uncomfortable with an act utilitarian recommendation in a harm-one-to-help-many scenario such as a trolley problem or forced organ donation, even if I know spillover societal impacts are contained, but I may still find that seeking to maximize utility feels like I am doing a good thing in most scenarios. I may believe that the Categorical Imperative has a lot of intuitive force to it while still being able to envision scenarios where lying would be a preferable action.

Yet still, I am motivated to seek out the ideal doctrine to follow in my pursuit of becoming a moral actor. I have faith that the centuries of work done by proponents of each system has a collective validity that far surpasses my gut reaction, but I cannot ignore these gut reactions entirely, nor can I simply select a system to follow arbitrarily and blindly abandon the rest. Additionally, I am drawn to the convenience factor of having an explicit set of rules to follow which I can treat as a “black box”--I want an ethical system which operates as a function for which I can use all relevant knowledge about some decision as input, and the output will tell me what action to take under that system. After all, my goal for being a moral actor is only worthwhile if it is actionable. I want to be able to go forth into the world and take good actions rather than spending my time pondering why and whether an ethical system is legitimate every time I wish to use it as guidance.

We are tasked with the problem of moral uncertainty\(^1\), clarified by Krister Bykvist in his paper of the same name, which in its most general form is simply the question of what to do “when we are uncertain about whether a certain moral theory is correct”. Bykvist explores several possible solutions to this problem, yet finds flaws and failure cases in each one, concluding that there is much work to still be done and that an ideal solution must be flexible
enough to account for large amounts of uncertainty. To achieve this goal, we need a mechanism with which we can evaluate ethical doctrines against each other with a sufficient degree of accuracy so as to learn more about their comparative benefits while being conscious of the pitfalls Bykvist brings to light. In this paper, I will propose such a system driven by experimentation and experiential data rather than pure argumentation to solve this, with the goal of converging on the ethical doctrine that most consistently recommends actions that I perceive to be good among the set of decisions I am faced with in my day-to-day life. The core of this system will be an implementation of an iterative, probabilistic algorithm called the multi-armed bandit, which is commonly used in research settings to converge on an optimal selection given a range of options. The procedure will involve engaging in consequential moral decisions over a period of time according to the guidance of a specific moral framework, and for each decision the actor will select the moral framework under which to act based on numerical information about all of the frameworks being considered derived from past iterations. Reflective intuitions about these actions will serve as our data for this experiment. I will argue why intuition about morality removed from a specific moral system is a legitimate source of data for evaluation and address some relevant concerns, and why a reflective approach is preferable. I will discuss how this system could be used in practice, what knowledge could be gained from such an experiment, practical limitations to the experiment, and possible modifications that allow us to forgo some of the assumptions that will be necessary for a simple version of the experiment.

One primary challenge of moral uncertainty, as Bykvist notes, is to determine what we “ought” to do at a level above any one moral theory, each with their own criteria for what we “ought” to do. A solution Bykvist puts forth is to suggest that for various moral theories we
determine the relative credences, or faiths we have in each one, treat those weights as likelihoods for the actions recommended, and take the action which seems to “dominate”. Yet it is not immediately clear why we would expect these numbers that we assign to our credences to align with their real probability of being, or why the action suggested by two moral theories instead of just one is more likely to be truly good. In fact, we cannot claim for certain that all moral rules are accurately expressed in some theory that has been discovered by human philosophers. If there is one optimal moral choice for any scenario, it may be the case that no one ethical system will identify every such choice correctly. Even if there is a true set of rules that if followed will always result in taking good actions, it may also be the case that we have yet to discover it. If it is possible that no present moral theory is accurate in all scenarios, it may be impossible to solve the dilemma of moral uncertainty to a degree of perfection, where we know for a fact that we are always making the true moral choice. Instead, our goal ought to be to do as best as we can, to seek out a set of rules that we can be confident is more accurate than our other available options. To achieve this task, we need some method of collecting data; we need to examine what actions are truly moral in a way that does not require the use of any one specific ethical theory. I will argue that human intuition is an ideal mechanism for this.

John Bengson\(^2\) defends a general intuitionist epistemology in “The Intellectual Given”, arguing that “intuition is… a presentation: a conscious state or event that, like perceptual experience, directly and immediately presents the world as being a certain way”, and that what he calls his quasi-perceptualist thesis -- that “intuitions are like perceptual experiences in being presentations” -- can be shown to grant epistemological weight to our intuitions. This follows from his claim that “having a presentation with the content \(p\) prima facie justifies one in
believing that $p$ if and only if one’s presentations are truth-conducive”. Essentially, if we believe that some presentation-receiving capacity of ours (such as visual perception) accurately isolates truth, we are justified in treating the objects of that presentation as knowledge. Andrew Sepielli\textsuperscript{3} takes a similar approach in developing a theory of rationality under moral uncertainty in “What To Do When You Don’t Know What to Do When You Don’t Know What to Do...”, developing concepts of perspectival (local) rationality and systemic (global) rationality to aid in constructing a probabilistic notion of normativity. In this sense, that which we receive as perspectival input can act as evidence pointing us in the direction of what we ought to do. I will argue can extend this epistemological framework to explicitly account for moral features in a way that lets us evaluate moral frameworks against each other. I claim that if we live in a world where morality is real and non-natural, for us to have any hope of being moral actors, moral intuitionism must hold at some level. Suppose otherwise; what would it mean if our gut intuitions about right and wrong had no bearing? At some point, every ethical theory must constructed either by someone who was unaware of a different ethical theory and “felt” some way about what is right, or by someone who “felt” that existing ethical theories were inadequate and in need of revision or replacement. When we evaluate arguments, we often form instinctive reactions about what they entail, which then go on to inform the types of logical responses we construct. It is thus justifiable to believe that the collective intuitions of mankind have played a prominent role in shaping our societies’ ethical theories over the course of centuries, and that if these theories have any legitimacy when considered as a whole, we must have done some valid work when using our intuitions to isolate abstract notions of goodness. For moral realism to hold, we must have some contact with whatever moral “stuff” is out there. I claim that this also makes sense from a
philosophy of mind perspective. Nearly every prominent ethical theory isolates some attribute of
the human mind as a justification for why we have moral worth or agency, such as capacity for
pleasure and pain, capacity to follow a moral system, capacity for autonomy, or consciousness. It
would then make sense for there to be some innate connection between the objectives of our
minds in isolating goodness and the true moral status of us and our actions. This is certainly not
to say that every intuition about morality has any accuracy to it; opponents of intuitionism⁴ often
point to the fact that genuine moral disagreement occurs between rational individuals to indicate
that intuitionism may be flawed. This implicitly assumes an all-or-nothing stance, where either
every intuition is valid or none of them are, which is unjustified and we need not accept in
defending intuitionism. Rather, in the pursuit of constructing a theory of intuition as moral data,
we merely need that our intuitions are correlated with true morality; they need only be right more
than they are wrong. If we wish to model some linear trend in the physical sciences, and we are
gathering data points which we will use to construct an equation that fits the model, our data
need not ever actually fall directly on the line. We can construct a line of best fit for the trend as
long as the data has some correlation better than a random distribution, and we can take a similar
approach with our imperfect intuitionist data. Of course, we would prefer that this correlation is
as strong as possible, but we can make do with simply knowing that there is a non-zero level of
legitimacy, which is granted to us by our intuitionist epistemology.

How then do we use our intuitions to evaluate various ethical frameworks? A starting
point for determining what types of moral thought are most accurate can be found in Peter
Railton’s “Moral Realism”.⁵ Railton’s character Lonnie, when sick and seeking some sort of
remedy while traveling in a foreign country, seeks out a glass of milk because it reminds him of
home. Yet, upon drinking the hard-to-digest milk, Lonnie’s stomach simply becomes more upset. Railton envisions an alternate character, Lonnie-Plus, who is a perfectly rational version of Lonnie with the power of hindsight to reflect on the scenario. Lonnie-Plus can determine that Lonnie’s homesickness led him to seek out milk, rather than a justifiable belief that it would cure his illness, and so Lonnie-Plus surmises that the “correct” thing to do in the scenario would be to seek out some sort of clear liquid. While we can’t hope to magically gain powers of hyperrationality like Lonnie-Plus, by taking the time to reflect on actions after we take them rather than forming judgments in the heat of the moment, we can hope to be more rigorous in our reasoning and more accurately isolate what the “correct” thing to do would have been. We have several reasons to believe that moral intuitions reflecting on actions we have previously taken are more valid. If we have yet to take an action, and there is one available choice that we may selfishly prefer but is likely not the most morally preferable, our intuition may become tainted by the desire for that choice and we can trick ourselves into thinking we are making a moral choice rather than a selfish choice. This biasing factor (like Lonnie’s homesick desire for milk) is drastically lessened after the fact, when the selfish goal has either been attained or forgone. Additionally, we simply have more information to work with in reflective evaluation, in the form of phenomenological data about the action we just took. The philosophical tradition of phenomenology, dating back to Edmund Husserl and his focus on treating vivid descriptions of his interactions with the world around him, is useful for constructing comprehensive intuitions about actions we have taken. We cannot intuit, beyond hypothesis or extrapolation from past similar experiences, how it will feel to take some action and perceive all of the relevant
phenomena. There may be aspects of an action that affect its moral standing that we are 
oblivious to until we take the action and are able to construct an intuition about it reflectively.

We want to make sure that we are not making too many assumptions when treating these 
reflective intuitions as a data source. Bengson, Terence Cuneo and Russ Shafer-Landau elucidate 
standards for what qualifies as legitimate metaethical data and methods in “Methods, Goals, and 
Data in Moral Philosophy”: under their view, metaethical data are “starting points for theoretical 
reflection” and they are “collected” from such sources as “intuition, introspection, common 
sense, ordinary experience, induction from experience, linguistic judgments, and observations”. 
We are safe in this respect. They state that data are neutral in the sense that they can be applied 
to several different theories, and that no individual data point should be considered infallible, but 
that the possibility that some data points may be “infused with” or “partial to” a particular theory 
is acceptable. By their analogy, “that physical data favor relativistic theory over Newtonian 
theory, for example, does not imply that the data unduly stack the deck”. We will keep these 
criteria in mind when making assumptions about reflective intuitions as data, but they satisfy 
these basic criteria and as such are freely usable as a data source. I will also use their notion of an 
ethical method as a “black box” of rules, where the output extends directly from the inputs, while developing the experiment that is central to this paper.

With reflective intuition as a legitimate source of moral data, we can begin constructing 
our experiment. The experimental model known as a multi-armed bandit is an ideal candidate for 
evaluating ethical doctrines first because it assumes very little about the inputs for evaluation and 
allows us to learn about their rates of success over time, but additionally because it includes a 
mathematical notion of “regret” in terms of mistakes it results in, which it successfully
minimizes. If we are willing to subject ourselves to an ethical experiment that will dictate decisions for us over the course of our day-to-day lives, we want a guarantee that we will not continue to use a certain ethical framework once we have determined with a high level of confidence that it is suboptimal. Multi-armed bandit experiments\(^8,9\) were initially conceived of as a mechanism to find the best slot machine (sometimes called a “one-armed bandit”) in a casino while losing as little money as possible. Now, they are commonly implemented in research to determine the most effective pharmaceutical drugs, optimal pricings for goods, and most user-friendly software features. The core motivation behind the experimental design is to find the ideal balance between “exploration”, where we are gaining information about the expected value of a particular function with little concern about the immediate payoff, with “exploitation”, where we choose a function that we think will result in the highest payoff. The initial input to a multi-armed bandit experiment consists of \(k\) “arms”, or functions, each of which has a set of possible payoffs with an unknown distribution. For our purposes, we can constrain this to functions with a binary output of 0 or 1. A function \(f_i\) would then output 1 with some initially unknown probability \(p_i\) and output 0 with probability \(1 - p\). The experiment takes the form of an iterative algorithm where each step relies on the resultant data from previous steps. At each step, we will choose one function from which to receive a payout, and we want to select the function that either best serves our purpose for exploration or exploitation.

To achieve this balancing act, we will maintain the notion of an Upper Confidence Bound for each function, which is essentially a guarantee that based on the past output data we have from the function, we can be confident with a certain probability \(1 - \delta\) that the expected value of the function is below some value \(x\). The expected value of a function’s output is a weighted
average of its outcomes and their probabilities, which for our function $\bar{f}$ would simply be:

$$1 \cdot p + 0 \cdot (1 - p) = p.$$ The Upper Confidence Bound for each of our $k$ functions will be initialized to 1, because we have no data indicating that the expected value must be smaller than 1. The computation for the Upper Confidence Bound requires the probabilistic notion of a bound on the sum of independent random variables called the Chernoff bound. I have attached a reference at the end of the paper detailing how it is calculated, but its specific mechanics are not crucial to understand for the sake of this paper. We can treat the Upper Confidence Bound as a black box that takes in past results as well as a desired confidence of $1 - \delta$ that outputs the correct upper bound. In practice, this Upper Confidence Bound is the ceiling of a confidence interval for the expected value of a function, and so as we learn more about a function by gathering data, we can be confident that the expected value will be in an increasingly tightening range. On any step where a function is called, we would expect the Upper Confidence Bound to decrease from its previous value unless the function returned a value greater than that previous value. On each iteration of the algorithm, we will choose the function with the largest Upper Confidence Bound to use, and in the case of a tie, select one of the maximal functions arbitrarily. Intuitively, the Upper Confidence Bound for a function will be large either in the case where we know relatively little about a function and still have a loose bound, or in the case where we have determined with reasonable confidence that a function’s expected value is fairly high, thus ensuring that we are either “exploring” or “exploiting”. In the initial iterations of a Multi-armed bandit, when we have little to no information, we expect to spend most of our time exploring. As our amount of information increases, we expect our frequency of exploitation to increase. Eventually we will expect our Upper Confidence Bounds to converge to the true expected values.
of each function, and so if there is one function with the maximum true expected value, we will select that function for every iteration in perpetuity.

We can define a mathematical notion of Regret, which indicates the difference between the sum of the outputs we receive in each iteration and the expected value of only choosing the true optimal function over the time it takes us to reach convergence. The Regret over \( T \) steps can be expressed as \( T \cdot x_{i^*} - \sum_{t=1}^{T} x_{i_t} \) where \( x_{i^*} \) is the expected value the maximum valued function and \( x_{i_t} \) is the value of the function selected on iteration \( t \). The Multi-armed bandit algorithm can be proven to result in a tight upper bound on the amount of Regret accrued over a period of iterations which improves as the period is extended. For a slot machine scenario this means that we have wasted as little money as possible in our pursuit of the most profitable machine, and for a drug trial scenario this would imply that we subjected as few patients as possible to ineffective drugs in order to determine the most effective one. When we apply the Multi-armed bandit model to our question of moral uncertainty, this means that we will take as few bad actions as possible over the course of the experiment before we reach convergence. We will explore a moral framework only as much as we need to in order to know that it is not worth pursuing indefinitely.

To adapt the Multi-armed bandit experiment to our comparison of ethical systems, if we have \( k \) different systems to consider, these will each be an “arm” represented as a function. Each iteration of the algorithm will be a scenario under which the experimenter must make a real-world decision of relative moral consequence. For decisions such as what to wear or what to eat for breakfast on a given day, there is likely not much value in the information gained if we try to evaluate these through a particular ethical system; most systems would not take strong stances
on these, and even if we could conjure a reason to believe that Kant would prefer oatmeal to toast, we would almost certainly not find any such action to be bad upon intuitive reflection. Thus we would be introducing positive feedback for that system into our dataset without really learning anything about whether it aligns with intuitive notions of goodness. because our intuitive reflection. Ideal scenarios are those which are common in day-to-day life but for which different doctrines may vary drastically in their prescriptions and for which we can have meaningful reflective intuitions. This may include decisions such as whether to give money to a homeless person, whether to donate to a charity that calls to ask for money, how or whether to vote in a local election, whether to violate a law banning littering or jaywalking, whether to buy meat or a meat substitute, or whether to spend money on a frivolous purchase. To carry out the experiment, we would formulate a list of such ideal scenarios, and also maintain a tabulation of each candidate system’s Upper Confidence Bounds, initialized to zero. We also would want some simple mechanism for resolving ties in Upper Confidence Bounds (such as repeatedly flipping a coin to generate a bit-string that would correspond to a number in some ordering of the tied systems). Whenever we encounter a scenario on our list while going about our day, we would choose the ethical system with the current highest Upper Confidence Bound and take the action in that scenario that the selected ethical system would recommend. After taking the action, we must then intuitively reflect on our perceptions of the action we just took to determine whether we believe we just did a good thing. If so, we tabulate a 1 as the result of that iteration and a 0 otherwise, and then recompute that system’s Upper Confidence Bound. We repeat this until we notice that we have been repeatedly selecting the same system for a substantial amount
of time, at which point we can accept this system as our ideal set of rules to follow in our pursuit of goodness.

This adaptation makes a number noteworthy assumptions about moral uncertainty. Some of these assumptions may be accounted for by further complicating the experimental model. Some result in a partial lessening of philosophical rigor, but I will argue that is not wholly detrimental to the goal of this experiment. First and foremost, because the Multi-armed bandit model seeks to maximize expected value, the goal of being a good person is best achieved when I am most consistently taking good actions. Essentially, we will have solved the problem of moral uncertainty if we have found a moral doctrine which we are certain will most consistently result in truly good action. Because we cannot know that any of our candidate theories is perfectly accurate, it is unclear how we could strive for something better than just finding the best. We could possibly devise some sort of weighting function such that decisions of greater consequence have a larger impact in our experiment, but this poses the challenge of fitting actions that we have qualitative opinions about to a quantitative scale. Additionally we assume that reflective moral intuitions take the form of a binary (1 or 0) response as to whether something is good or bad. This may be imprecise, as some intuitions about goodness could have much more perceptual force than others; however, because it is difficult to assign a precise quantitative ranking as to how strongly we feel an intuition and much easier to qualitatively determine whether we have that intuition, a scaled system that maintains practical accuracy may be near-impossible. Over time, we would expect the frequency of situations where we feel a particular strength of intuition when acting under different ethical systems to balance out, and thus this is unlikely to affect the outcome of our experiment. Because we can constrain the list of
candidate scenarios at the beginning of the experiment, we can choose a list that allows us to be fairly confident that no one decision will have drastically more moral gravity than some other. Additionally, a binary system solves issues that come up with deontological systems, which Bykvist tells us to take note of. A deontological system may hold that any action which is not strictly good is strictly bad, and thus a binary evaluation is the best we can do if we wish to allow our intuitions to converge on a such a system. We also assume that the outcome of each decision is probabilistically independent, meaning that whether the choice we make in iteration $i$ of the experiment is good has no direct bearing on whether the choice we make in iteration $j$ is good. This is a mathematical requirement of the calculation of the Chernoff bound, and should not be too difficult to enforce if we are conscious to omit from consideration scenarios in which it seems apparent that our expected intuition about some action may be directly affected by an intuition regarding a previous action (such as if we find that we’ve been dwelling on an intuition for). If we were to forgo this assumption of independence, there are other formulas (Markov’s Inequality, Chebyshev’s Inequality) which we can use to calculate an upper bound, but they do not yield nearly as tight of an estimate and thus would take more iterations to reach convergence. While these assumptions are important to be aware of

We now have an experiment that allows us to seek an answer to the dilemma of moral uncertainty with minimal regret and a sufficient degree of rigor. If we were to devote a considerable amount of time into clarifying a comprehensive set of ethical doctrines and test them out in our daily routines, we could strive for moral convergence via our Multi-armed bandit model. If we were to enlist a swarm of willing participants who could gather data in parallel, we could conduct the experiment more efficiently and potentially But ultimately, the aim of this
paper is not to convince the reader to spend months or years of their life forcefully making decisions under the recommendation of one of an explicit list of ethical theories which they select by tabulating all of their past intuitions in a notebook in their pocket. It is to elucidate, through the allegory of a rigorous research experiment, that when convinced to pursue morality yet faced with the dilemma of moral uncertainty, we can use our lived experiences to seek legitimate answers. Even the most devoted scholars of act utilitarianism, Kantianism or Christianity are likely to stray away from the recommendations of their respective systems in their daily lives. We have a finite capacity of energy and time that we can use to compute utility distributions of various actions or consider how the Categorical Imperative implies to actions we are faced with for the first time. I argue that if this divergence from a pure theory is inevitable, possibly in part because there are scenarios where our intuitions about how to act draw us away from a doctrine that we subscribe to, it is worth approaching moral uncertainty from the perspective of what is consistently actionable. It is possible to trust in the effectiveness of a moral system without needing to be able to argue directly why it is superior to all others. If we experience the success of one system through our own phenomenological and intuitive perception, and experience the failure of others, we can become convinced of that system’s probable truth without having to sort through the years of literature arguing for and against it. By ignoring a purely ethical pursuit of the most argumentatively robust theory, we can focus on a more practical philosophical goal: learning how to make choices such that we maximize the amount of good action we take. It is worth noting that if every student of philosophy genuinely attempted this experiment, beginning with the same well-defined lists of candidate theories and candidate scenarios, we would not expect everyone to converge on the same moral framework.
Some of this would be due to differences in intuition as well as experimental randomness, but even accounting for those, there is still another important factor at play. If no moral theory we currently have isolates the correct moral action in every scenario, then every theory will have a distribution of accuracy over the set of all possible actions we as humans can take. Every person will engage in a distinct subset of those actions over the course of their lifetime, and thus the theory which results in some Person A taking the highest frequency of moral action may not have the same result for some Person B. From a practical perspective, if we wish to encourage people to take morally good actions, it might be most advantageous to not recommend the same set of rules to everyone. Determining what those rules precisely are for everyone is a daunting task, but our approach gives us a starting point and a blueprint for developing other theories in the future. Moral uncertainty is unique among many meta-ethical dilemmas in that answers to it are prescriptive in the sense that they recommend specific courses of action; uncertainty ventures down into the realm of pure ethics while also requiring interaction with abstract notions of goodness removed from a specific ethical doctrine. This actionability alone should motivate a considerable amount of additional exploration into the topic. Bykvist notes¹ that resolving the question of moral uncertainty is of pressing importance for addressing generational and epistemic concerns such as global warming. Because the nature of this experiment requires a fairly large volume of data, forcing us to consider common, everyday questions about moral action as our candidate scenarios, we might not expect the set of rules we converge upon to maintain their accuracy for situations of far greater moral concern. But by generally adopting an approach of giving credence to our reflective moral intuitions, and using them as data to inform our future decisions, we can get a little bit closer.
References


Chernoff/UCB Computation and Proof of Regret Minimization (Roth):

http://www.cis.upenn.edu/~aaroth/courses/slides/agt17/lect20.pdf