Moral Life Under Phenomenological Uncertainty

Vladmir Chituc

I wrote this prospectus in the third year of my program (and within the first few months of the COVID-19 pandemic), so it's a weird document in a lot of ways (and weird to look back on, now). Most immediately, literally none of this actually made it into my dissertation — this was the earliest version of a lot of those ideas, and my initial plan involved a lot more in person, lab-based experiments. As COVID dragged on, I did a lot more online experiments, and the few online studies I sketched here didn't really pan out, so I just had not much else to do in terms of research except read old papers and think through some of the more complicated theoretical issues. Eventually after maybe a year and half, things really clicked and I ended up developing an entirely new research program based on some of this (often overlooked) older work, and my research went in an ultimately cooler (I think) but still related direction. So my dissertation is almost entirely work I did in the last year or so of my PhD. But for the prospectus, the overall ideas and structure was just something I did in concert with my advisor over a few months of meetings. We'd discussed a lot of these ideas already and were thinking about each chapter more or less as an empirical project that was primarily going to turn into a publication (and would then be turned into a dissertation chapter). Once the deadline was getting closer, I presented a broader overview to the lab to get their feedback, which helped flesh out some of the more nitty gritty details.

In terms of the actual writing of it: I wrote the long version first, and I had a little bit of trouble shortening it for the short version (but the advice I was given is that only the long one really matters in terms of being evaluated, the short one is basically just a rough overview to send to the university so that there's something on file; I sincerely doubt that anyone ever read it). In general, though, I was often given the advice to just have your dissertation be three empirical papers you've stapled together with a few dozen pages as intro and discussion, and I think people were encouraged to more or less model their prospectuses on that. That's not what I ultimately ended up doing, but I owe that largely to the fact that all of the work I wrote up was done within a few months of each other and I wrote them all up around the same time, so there's a lot more coherence than I think there otherwise would have been. And like everything else I've ever written, I had everything conceptually figured out well in advance but didn't start actually writing any of it up until a few days before the initial deadline.

Introduction (long version)

Some people like the taste of cilantro, and others find that it tastes like dish soap. That two people may experience the taste of an herb differently is nearly too trivial to note, but the implications of this fact are not. Consider a straightforward example. Two friends are at dinner, and they are deciding on an appetizer to share: guacamole or salsa. One friend prefers salsa to guacamole, liking the spiciness of salsa and finding the cilantro in the guacamole to have a soapy taste. The other friend has the opposite preference, liking cilantro and finding salsa too spicy. Assuming the two friends must choose a single dish to share, and assuming each gives equal weight to their own wellbeing and the wellbeing of their friend, then how are they to decide which dish would be best for them, overall?

This mundane example illustrates a commonplace and underappreciated aspect of decision-making: we must often make decisions about an outcome we haven't experienced and cannot imagine. Neither friend can know how either dish will taste to the other, and this is especially clear for the friend who dislikes cilantro. While the friend who likes it might crudely imagine adding a bit of soap to guacamole, it's not clear how one might mentally subtract soapy flavor to taste the herb as everyone else does. Without such knowledge, however, it's not clear how either friend can accurately represent the value each option holds overall. Economists and philosophers call decisions like this *interpersonal utility comparisons* (for early discussions, see Hammond, 1977; Robbins, 1935, ch 6; Suppes & Winet, 1955), and they have long posed a puzzle, with some scholars going so far as to argue that such comparisons are impossible (see e.g. Hausman, 1995).

Whether these decisions are impossible or not, we nonetheless make them constantly and without too much thought in our everyday lives. In my dissertation work, I argue that these decisions pose a problem because they involve what we call *phenomenological uncertainty* (Chituc, Paul, & Crockett, In Prep). Put another way, these are decisions where we are unsure of what it is subjectively like to experience a relevant outcome. Here, we ask how this uncertainty is resolved, how it differs from other kinds of uncertainty, and how we might resolve it better. It may not be troubling that friends may make suboptimal decisions about what appetizer to order, but there are many such decisions with larger stakes. Civilians decide whether to send troops to war, largely male legislatures write laws about reproductive health, and university administrators set policy for students with drastically different backgrounds, identities, and life experiences.

In my dissertation, I aim to explore phenomenological uncertainty using a number of approaches. First, I present an overview of the theoretical and empirical background of this work, including a review of empirical work on decision-making under uncertainty. Phenomenological uncertainty is most closely related to impact uncertainty (Kappes et al., 2018), and I will pay particular attention to this relationship. Next, I discuss so-called "supertasters," as well as relevant work in psychophysics, which serves as the basis for our paradigm. Then, I present an overview of the proposed studies, including predictions and their rationales. Finally, I discuss the implications of this work, my proposed timeline, and my contingency plans given the ongoing COVID-19 pandemic.

Decision-making under uncertainty

Far from being a unified concept, there are a number of varieties of uncertainty (Kappes et al., 2019), and these different kinds of uncertainty have different profiles and

downstream consequences for decision-making. Consider two kinds of uncertainty: uncertainty around the likelihood of an outcome, or *outcome uncertainty* (e.g. Platt & Huettel, 2008), and uncertainty about how that outcome might affect other people, or *impact uncertainty* (Kappes et al., 2018).

When playing a dictator game, outcome uncertainty lead to decisions that were less prosocial. When payoffs for the other person were hidden, participants chose more selfish outcomes (Dana et al., 2007). In another study, Kappes and colleagues (2018) replicated the effect of outcome uncertainty on decisions in a dictator game, but they contrasted those results with the effects of impact uncertainty. Instead of introducing uncertainty about whether or not an outcome would be fair or not, they learned how big an effect the outcome would have on the other person. In this study, participants also played a dictator game, but they sometimes learned about the socio-economic status of the receiver. Sometimes the receiver was poor, rich, or, in a case involving impact uncertainty, somewhere along the distribution from rich to poor. Compared to a standard dictator game, which included no information about the recipient, participants were most prosocial in both the certainly poor and impact uncertainty conditions, suggesting that one strategy for decision-making uncertainty is precautionary, assuming the worst-case scenario.

Of the kinds of uncertainty described above, phenomenological uncertainty seems most closely related to impact uncertainty. It may be the case that the subjective character of an experience operates mostly along the same lines as something like socioeconomic status. Just as some people are more or less impacted by gaining or losing money, depending on whether they are poor or rich, other people might be more or less impacted

by eating cilantro, depending on whether it tastes soapy to them or not. Even so, it seems as if this is not the whole picture—we might know that someone dislikes cilantro while nonetheless remaining uncertain about what it's actually *like* for that person to taste it.

Thus, we propose in our theoretical model (Chituc, Paul, & Crockett, In Prep) that impact and phenomenological uncertainty are distinct. To briefly summarize our view, one way we assign an outcome value is through reference to experience. This might be retrospective, in that we might retrieve a value that we have updated over time or remember specific instances of that experiences (e.g. Bornstein et al., 2017; Gershman & Daw, 2017; Murty et al., 2016). It might also be prospective, in that we might simulate novel experiences based on relevant past experiences (Barron et al., 2013). Taken together, this suggests that resolving phenomenological uncertainty should resolve impact uncertainty, but not vice versa. Put another way, knowing what an experience is like allows us to assign that experience value, but knowing the value of an experience does not tell us what it's like.

In order to study impact and phenomenological uncertainty in the lab, we need a paradigm that allows us to manipulate these factors experimentally. To do this, a stimulus must meet a few criteria. First, this stimulus must be experienced in different ways by different people, such that only some are able to know what the experience is subjectively like. Second, this stimulus must be presented in a way that minimizes impact uncertainty. Put another way, while only some people should know what the stimulus is subjectively like, everyone should know how good or bad that experience is. Finally, it must be practically and ethically possible to administer this stimulus in a lab setting.

To meet these criteria, we make use of a phenomenon identified in the psychophysics literature—some people are able to taste phenylthiocarbamide (PTC) and a related compound, 6-n-propylthiouracil (PROP), while others are not. Nearly 30% of Caucasian American adults perceive no taste to PTC or PROP, and the remaining 70% of Americans perceive it as more or less intensely bitter, depending on whether they are so-called "supertasters" or "mediumtasters" (Tepper, 1998). Except when this distinction is relevant, I will refer to supertasters and mediumtasters as "tasters" for simplicity.

Importantly, tasters and nontasters don't only differ in whether they are able to taste certain compounds. Rather, tasters experience flavors more intensely, in general, and they have a higher density of taste buds on their tongue (Bartoshuk et al., 1994). This discovery was made only recently, however, exactly because no one has direct access to how anyone else experiences taste. Consider that one method for categorizing tasters explicitly operated on the assumption that tasters and nontasters did not differ at all in how intensely they experienced salty flavors, thus underestimating the differences between the two groups (see Bartoshuk et al., 2004, p111). Accordingly, the standard tools of psychophysics distorted our understanding of tasters and nontasters for decades.

As a simple illustration of this basic problem, consider how we might measure how two people experience the bitterness of a lemon. Though tasters and nontasters may have very different experiences, they may nonetheless provide identical ratings using a 10-point scale that ranged from "not at all bitter" to "very bitter." This is because adjectives like "very " are relative, and a lemon would indeed be very bitter to both the taster and nontaster when considering the range of bitterness each has experienced (c.f.

Birnbaum, 1999). Thus, we might miss a true difference between tasters and nontasters, believing that both experience a stimulus as equally bitter when in reality they don't.

Even more, such scales might do more than hide differences that exist; they might suggest that one person experiences something more intensely when the opposite is true. It's possible that a taster may experience something "moderately bitter," like black coffee, as objectively more intense than a nontaster might experience something "very bitter," like a lemon (see Fig. 1 below). This is because the taster has a wider range of intensity flattened into what seems like a directly comparable set of adjectives. This basic problem has sparked extensive work in psychophysics that has endeavored to produce measures that might compare objective intensity of taste and other experiences across groups of people (for reviews, see: Bartoshuk & Snyder, 2004; Snyder, Sims, & Bartoshuk, 2015).

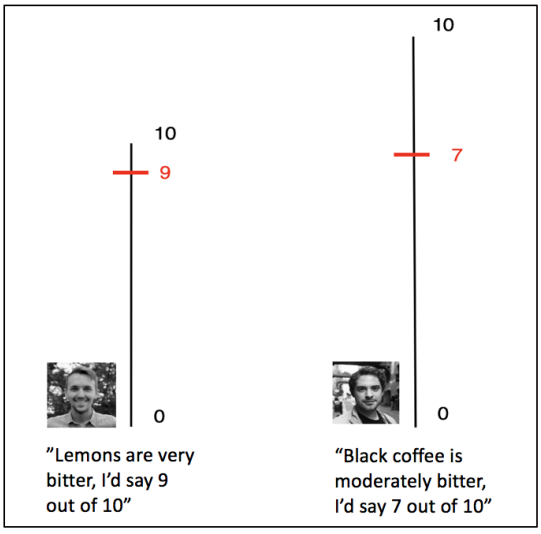

Fig. 1. A simple illustration of how scales can be misleading given individual differences in taste sensitivity (see also Prutkin et al., 2000). Some people, like so-called "supertasters" (right), experience tastes more intensely than others do (left). Thus, using a common scale can lead to scenarios where someone providing a lower rating might nonetheless experience something more intensely.

To address this problem, research in psychophysics has stressed that there must be some common point of reference if we hope to meaningfully compare one person's response on a scale to another's (see e.g. Bartoshuk et al., 1994; Marks et al., 1988). Thus, researchers in this area have conducted extensive work aiming to measure how someone's taste experiences might differ according to their PROP sensitivity (for an overview, see: Bartoshuk & Snyder, 2004; Snyder et al., 2015).

Early work used labeled scales to classify taster status, often comparing the intensity of different concentrations of PROP with different concentrations of salt solution. This work suffered from the problem described above, however, underestimating the difference between tasters and nontasters. For example, one study (Drewnowski et al., 1997) found that intensity ratings could distinguish nontasters from tasters, but this scale found no differences in perceptions of saltiness across taster status. Findings like this contributed to the false impression that PROP sensitivity was independent of perceptions of other flavors, like saltiness or sweetness.

To avoid this problem, research in psychophysics aimed to develop a scale that could more directly compare experiences by referencing an independent modality. Put simply, the logic behind this strategy is that, on average, it's unlikely that tasters and nontasters differ on more than their sensitivity to tastes, and they should experience other modalities like sound, light, pain, and so on, in comparable ways. Thus, researchers developed the generalized labeled magnitude scale (gLMS), a quasi-logarithmic scale that ranges from 0 (no sensation) to 100 (strongest imaginable sensation of any kind). So long as the strongest imaginable sensation of any kind is not correlated with taste, and so

long as the groups being compared don't systematically vary along the strength of sensations imaginable, then this scale can be used to rate the intensity of any kind of experience across any modality (Bartoshuk et al., 2004).

Taster status and phenomenological uncertainty

Earlier, I outlined several different kinds of uncertainty—outcome uncertainty, impact uncertainty, and phenomenological uncertainty. Based on the discussion above, consider the uncertainty involved when a nontaster considers the experience of a taster who has sampled a solution of PROP. It's unlikely that outcome uncertainty is at play, since the nontaster can know that the solution contains PROP and not water. Impact uncertainty is less straightforward, however. While impact uncertainty is in some ways minimized, since the nontaster knows that the solution tastes bitter to the taster, the nontaster has no clear way to know just how intensely bitter the solution is, at least in most cases. This kind of uncertainty is not unresolvable, however, especially in light of the work in psychophysics reviewed above. If a taster rated the PROP solution on the gLMS described above, and if this rating was shared with the nontaster, then the nontaster would have a point of reference through which they might resolve impact uncertainty. Thus, the nontaster might compare the impact of PROP to similarly intense experiences, be it loud noises, bright lights, remembered pain, and so on. Thus, we might resolve both outcome and impact uncertainty, but it's not clear that we are any closer to resolving phenomenological uncertainty. Put another way, it's not clear that the nontaster can know what PROP is actually *like* for the taster, and it's not clear how a nontaster might ever know.

In sum, when considering the experience of a taster sampling PROP, nontasters are in a position of both impact and phenomenological uncertainty. While it seems possible to resolve impact uncertainty through information we might learn from scales like the gLMS, it's not clear that it's possible to resolve phenomenological uncertainty. Thus, phenomenological uncertainty seems apart from other kinds of uncertainty discussed in the literature, though my dissertation work aims to test this empirically.

If phenomenological uncertainty is a distinct kind of uncertainty, how, then, do we make decisions that involve it? Here, I consider two possibilities, simulation and precautionary preferences. I discuss them in that order.

Simulation and its pitfalls

In the literature on interpersonal utility comparisons, there is one solution to this problem worth giving particular attention. This solution is grounded in *empathic projection,* sometimes called *extended sympathy* (see e.g. Sen, 1974), and it is closely related to modern literatures on empathic accuracy (Ickes et al., 1990), social projection (Krueger & Clement, 1997), and simulation (Shanton & Goldman, 2010).

Despite the extensive writings of philosophers and economists, people nonetheless find little trouble making interpersonal utility comparisons. In explaining this, Harsanyi (1977) suggests the following:

> *Simple reflection will show that the basic intellectual operation in such interpersonal utility comparisons is imaginative empathy. We imagine ourselves to be in the shoes of another person, and ask ourselves the question, "If I were now really in* his *position, and had* his *taste,* his *education,* his *social background,* his *cultural values, and* his *psychological make-up, then what would now be my preferences between various alternatives, and how much satisfaction or dissatisfaction would I derive from any given alternative?" (p. 638)*

Though this may be true for how phenomenological uncertainty is actually resolved, and thus how interpersonal utility comparisons are actually made, we need one more step for this to be accurate or useful. Harsanyi calls this *the similarity postulate.* If we estimate how good an outcome will be for someone else by simulating how good it would be for ourselves, we must also assume that everyone involved experiences things in basically the same way. Then, once granted some of the caveats above, such as differences in taste, education, and so on, Harsanyi writes "it is reasonable for me to assume that our basic psychological reactions to any given alternative will be otherwise much the same (p. 639)." Some, such as Goldman (1995), consider this to essentially solve the problem of interpersonal utility comparisons.

Nonetheless, Harsanyi cautions that we may misapply the similarity postulate. He writes: "For instance, I may fail to make proper allowances for differences in our tastes, and may try to judge the satisfaction that a devoted fish eater derives from eating fish in terms of my own intense dislike for any kind of sea food." He goes on to reassure us, however. "Of course, sensible people will seldom make such an obvious mistake (p. 639)."

Others have argued, however, that the solution to this problem is not so simple. For example, Adler (2014) claims that such empathic projection fails when these comparisons involve traits that are unique to an identity, impossible to simulate, or are relevant to wellbeing. Even more, it's not clear how we are to handle cases where the similarity postulate plainly does not hold, as is the case with a nontaster considering the experience of a taster. Consider, too, the literature on egocentric biases. We are often prone to judge others to be more like ourselves than they really are (Ross, Greene, &

House, 1977), and our own preferences often interfere with our ability to learn the preferences of others (Gilovich, 1990; Tarantola et al., 2017). Furthermore, our own experiences affect how we anticipate the experiences of others. For example, if we are bored of a joke, we expect others who hear it for the first time to find it less funny, too (Campbell et al., 2014). Though this is only a small sample of relevant work in this area, this suggests that, if we resolve phenomenological uncertainty through simulation, we are likely to misjudge the experience of another person as overly similar to our own.

Harsanyi's proposal is nonetheless compelling for a number of reasons. First, it seems to be a plausible and intuitive account for how we resolve phenomenological uncertainty, and in doing so, it explains how we are able to constantly and without much effort make decisions that some philosophers and economists consider impossible. Second, it makes a clear prediction for how decisions under phenomenal uncertainty may be lead astray, such as in cases where the similarity postulate doesn't hold or when empathic projection fails.

Precautionary preferences

Though we argue that phenomenological uncertainty is distinct from impact uncertainty, research on the latter may provide insights into how we decide under phenomenological uncertainty. To return to work that compares impact and outcome uncertainty (Kappes et al., 2018), there were two conditions where participants were most prosocial in a dictator game: when the recipient was poor, and when it was uncertain whether the recipient was poor or well off. This is explained by *precautionary preferences,* where we tend to make decisions as if the worst outcome were true given certain kinds of uncertainty (Crockett et al., 2014; Kappes et al., 2018). This strategy is

called “the precautionary principle” in public policy, in that it advocates caution when introducing a new technology or policy until it’s proven safe (Sunstein, 2005).

Thus, in the *impact uncertainty* condition described above, participants may have been making decisions using a heuristic like “assume the recipient is poor.” Accordingly, when making a decision that requires us to resolve phenomenological uncertainty, we may simply leave that uncertainty unresolved and follow a basic heuristic like “avoid bad outcomes” or “assume the worst.”

Overview of Proposed Studies

We can raise a number of open questions from this brief literature review. How do we decide under phenomenological uncertainty, given that we don’t know what it’s like to experience a relevant outcome? Do we use simulation or empathic projection, trying to imagine what the experience would be like for us? Or do we err on the side of caution, assuming the worst and operating under something like the precautionary principle? We might also ask to what extent phenomenological uncertainty is distinct from related concepts, like impact uncertainty. Finally, we might consider the kinds of moral inferences we make when people decide under phenomenological uncertainty. If someone inflicts a harm that they don’t experience as harmful, do we hold them less responsible?

To answer the questions described above, I propose a number of studies. I start by describing our harm aversion task, as well as our computational model of moral decision-making, both of which are adapted from Crockett and colleagues (2014) and validated across a number of behavioral and neuroimaging studies (e.g. Crockett et al., 2017;

Siegel et al., 2018). I start by discussing pilot work and a hypothetical version of this task (Study 1), as well as an in-person version of the task (Studies 2a & 2b).

In a second set of studies, we explore the role of testimony in making decisions under phenomenological uncertainty (Study 3). Since we lack direct access to the experience of another person, when we make decisions that affect other people we must often consider what a person says their experience is like. Here, we ask whether some forms of testimony are more persuasive than others, and in doing so, we are able to explore to what extent impact and phenomenological uncertainty are distinct concepts.

Finally, I explore the moral judgments and inferences we make when others decide under phenomenological uncertainty. Put simply: who do we judge more harshly for inflicting harm, someone who has experienced that harm and knows what it's like, or someone who hasn't? We aim to start testing this idea with a vignette study (Study 4).

Our stimuli and scale

As a stimulus that some are able to experience and others aren't, we will use five different concentrations of 6-n-propylthiouracil (PROP), 0.032, 0.18, 0.32, 1.8, and 3.2 mmol/L. These concentrations have been used across a number of studies to classify participants according to genetic PROP taste sensitivity, though some tests use only one or three (e.g. Ditschun & Guinard, 2004; Tepper et al., 2001). Each sample will be delivered to participants in 0.2 ml micropipettes, and participants will rinse their mouths with spring water before and after each sample.

To classify participants according to taster status, they will rate each sample using the gLMS, described above (Bartoshuk et al., 2004). This scale ranges from "no sensation" (0) to "strongest imaginable sensation of any kind" (100), with adjectives

distributed quasi-logarithmically throughout, including barely detectable (1.6), weak (6), moderate (17), strong (35), and very strong (51). To orient participants to the scale, participants will first rate 15 imagined or remembered sensations, followed by five 1000-Hz tones (Duffy et al., 2004; Hayes et al., 2013). In line with past work (Tepper et al., 2001), we will classify participants as "tasters" if they rate the most concentrated solution higher than a 15 using this scale. When it's relevant to further subdivide tasters, we will classify supertasters as those rating higher than 51 on the scale, and medium tasters those who rated below 51 on the scale.

Our task and model

There are two roles in this task, a "Decider" and a "Receiver." Two participants will arrive separately and remain anonymous to one another. The Decider will make a series of choices between higher quantities of PROP solution for more money and lower quantities of PROP solution for less money. While the Decider always receives the money, the PROP solution is administered to the Decider in only half of the trials ("Self" condition). In the other half of trials, the PROP solution is administered to the Receiver ("Other" condition). One trial will be randomly selected at the end of the experiment, then implemented.

The trials will involve between 1 and 20 doses of the .2 mL micropipettes used to determine taster status earlier in the experiment, and the monetary incentives will range from $0.10 to $19.90.

The model we use in the task is as follows:

$$\Delta V = (1 - \kappa)\Delta m - \kappa \Delta s$$

$$\kappa = \begin{cases} \kappa_{\text{self}} & \text{if Self Trial} \\ \kappa_{\text{other}} & \text{if Other Trial} \end{cases}$$

Here, ΔV is the difference in subjective value between the two options, while Δm and Δs are the objective differences between the amount of money and amount of PROP solution, respectively. Harm-aversion is captured by the parameter κ, which is between 0 and 1. If a participant is maximally harm averse, they will forgo any amount of money in order to deliver the smallest possible amount of solution, and their κ will be 1. If a participant is minimally harm averse, then they will always administer any amount of PROP solution to receive more money, and their κ will be 0. Importantly, harm aversion is modeled separately for when the Decider chooses for themselves (κ_{self}) or for the Receiver (κ_{other}).

As described above, while the Decider in this task will sometimes be a taster and other times not, the apparent Receiver in this task will always be a taster, which the Decider will be aware of. This will allow us to compare κ_{other} for both tasters and nontasters, which will allow us to see how harm aversion varies based on whether phenomenological uncertainty is present (when nontasters decide for tasters) or absent (when tasters decide for tasters). To parse phenomenological uncertainty from impact uncertainty, participants will always see the rating the taster provided using the gLMS. Thus, all participants will know how intensely bitter the PROP solution is, but only some will know what the experience is subjectively like.

Predictions and rationale

In this first set of studies, I aim to explore how harm aversion might vary based on phenomenological uncertainty. In doing so, I test two competing hypotheses. If decisions like this are made through a process of empathetic projection, such that the decider simulates what an outcome would be like were they to experience it, then an

independent t-test should reveal an egocentric bias, such that nontasters are less harm averse than tasters (have a κ_{other} closer to 0). This is because their own experience with PROP suggests it is only mildly bitter, if not completely neutral. Furthermore, empathic projection should predict a linear relationship between κ_{other} and the participants own rating of the intensity of PROP. Using a simple linear regression, we would expect that for both tasters and nontasters, their own ratings of the PROP solution should predict κ_{other}, such that those with lower ratings should have κ_{other} closer to 0 and those with higher ratings should have κ_{other} closer to 1.

In contrast, precautionary preferences, as developed in work on impact uncertainty (Kappes et al., 2018), would predict that an independent t-test should reveal that nontasters should be equally if not more harm averse than tasters (have a κ_{other} closer to 1). Thus, nontasters would be assuming the worst and making decisions accordingly. Furthermore, we would expect that a regression using ratings of PROP solution should only predict κ_{other} for tasters, since the precautionary preference account does not depend on one's own experience.

Study 1: A hypothetical task

To use PROP solution as a stimulus for this task, we must consider a few constraints. First, in the harm aversion task this paradigm is adapted from (Crockett et al., 2014), the Decider or Receiver could experience up to 20 shocks. Thus, PROP must be administered in doses small enough that it can be administered many times over.

A second constraint is that existing work aiming to classify participants according to PROP sensitivity (e.g. Tepper et al., 2001) used 10mL samples of PROP solution. In a

small pilot (n = 3), the author and his collaborators sampled a small quantity of each PROP solution (less than 1ml each), and this revealed a few important details. First, all three participants were supertasters (rating the 3.2 mmol/L solution over a 95 on the gLMS), and second, all agreed that it would be unreasonable to ask participants to sample 10mL of this solution 20 times (Crockett, M.J & Paul, L.A, personal communication, 2019; Chituc, V., personal experience, 2019).

Thus, we conducted a pilot with 55 participants recruited from the Yale Introduction to Psychology Subject Pool. As described above, participants oriented to the scale using imagined and remembered sensations, then sounds. Next, we administered PROP solution using 0.2mL micropipettes, with the aim of determining whether this small quantity of solution could produce comparable ratings on the gLMS as larger volumes of PROP. To more directly compare these ratings with past work, we divide tasters into supertasters and mediumtasters. The results are plotted in Fig. 2 below.

Using the cutoffs for taster status described above, we found that 18% of our sample were supertasters (ratings above 51 on the gLMS), 55% were mediumtasters (ratings above 15 on the gLMS), and 27% were nontasters. As a point of comparison, Tepper and colleagues (2001) found in their sample 18% supertasters, 57% mediumtasters, and 25% nontasters. A repeated measures ANOVA revealed that intensity ratings of PROP significantly differed across taster status $F(2,52) = 80.04$, $p < .001$. Thus, we are confident that 0.2mL concentrations of PROP can both reliably categorize participants by taster status and serve as stimuli in our harm aversion task.

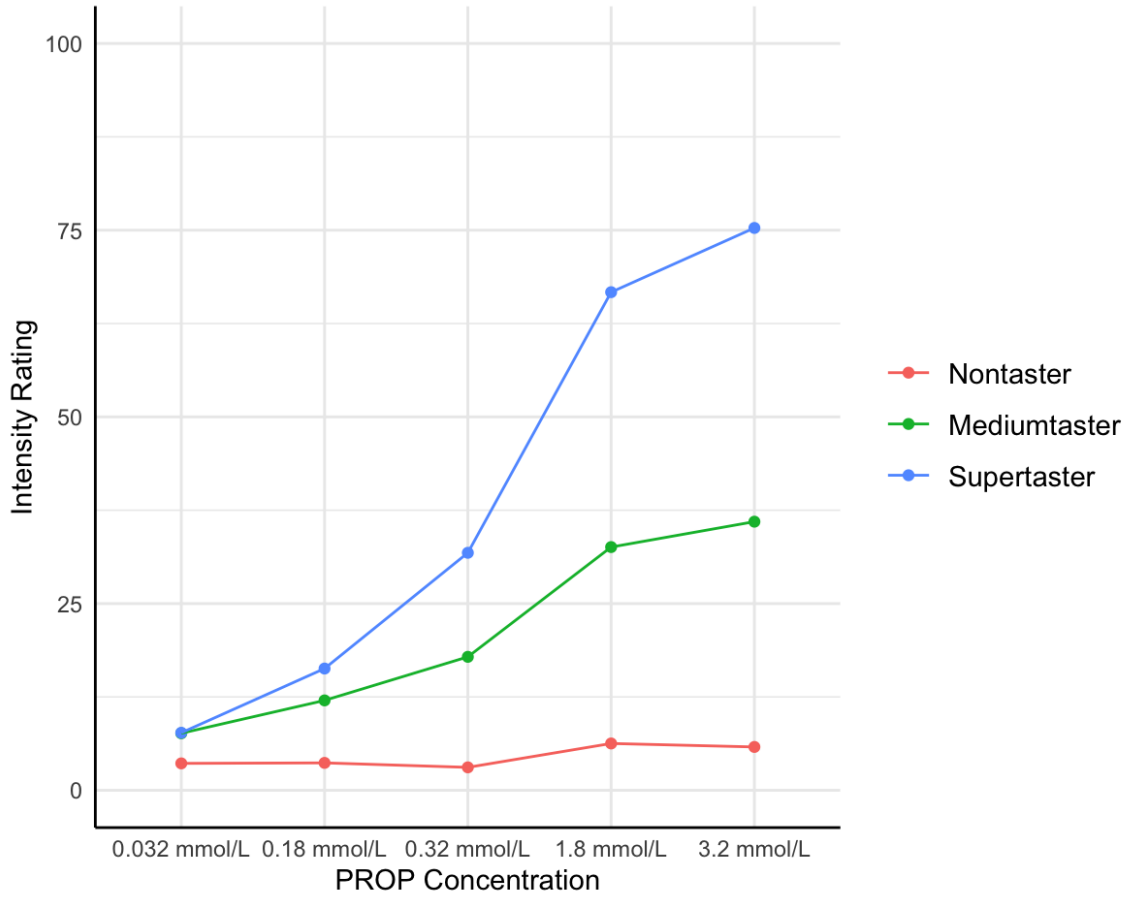

Fig. 2. Mean intensity ratings of different quantities of PROP solution across taster status. Stimuli were rated using the generalized Linear Magnitude Scale.

Of the 55 participants who took part in our pilot, 31 completed a hypothetical version of our main task as described above. After sampling and rating the PROP solutions, participants made 20 decisions between different amounts of money and different numbers of doses of PROP solution to be delivered to a hypothetical Receiver. Participants read that the dose of PROP solution would be one 0.2mL sample of the most concentrated solution they tasted earlier, and that they would receive the money in this task while another participant would receive the doses of PROP. Put another way, these were all "Other" trials of the task. Finally, participants read that the solution they sampled tasted neutral to some and bitter to others, and that the hypothetical receiver in the task

rated the solution as an 81 on the gLMS (this was the mean rating provided by supertasters thus far in piloting).

Since both supertasters and mediumtasters experience PROP as bitter, we categorize them both as tasters in this analysis. Of the 31 participants who completed the task, 8 were nontasters and 23 were tasters. After calculating κ_{other} for each participant, no clear pattern between taster status or individual rating of the PROP solution has emerged thus far (see Fig 3. below). The ongoing COVID crisis interrupted piloting of this task, and it will continue once we are able to safely conduct human subjects research in the lab.

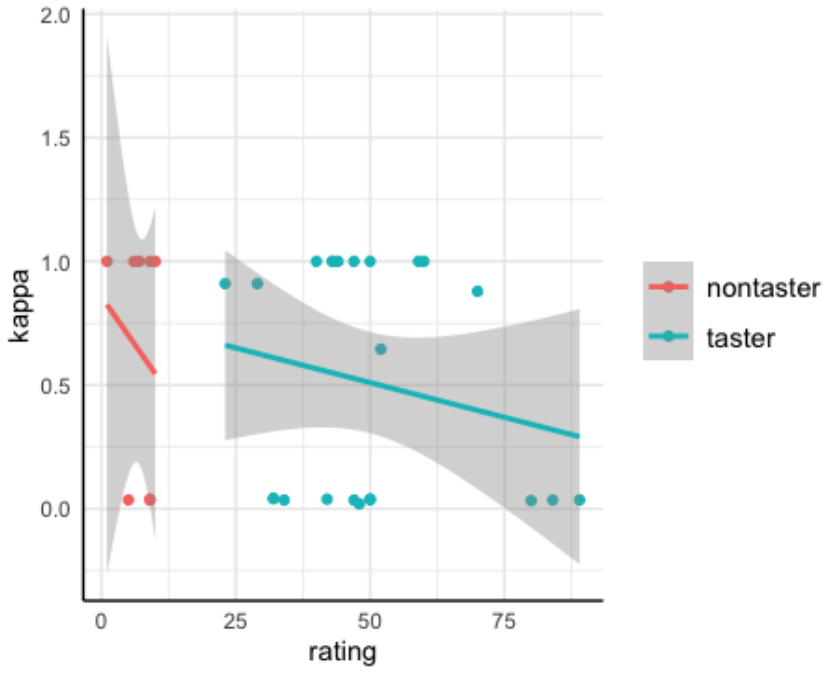

Fig. 3. Harm aversion (kappa) in a hypothetical task based on taster status, plotted against the rating each participant gave for the PROP solution. Higher kappa values represent a higher aversion to harm.

This first study will inform our future work in a number of ways. First, it allows us to begin to test how we decide, or at least anticipate deciding, under phenomenological

uncertainty. If we decide through empathic projection, we should expect a t-test to reveal that tasters are more harm averse than nontasters, and that there is a linear relationship such that kappa becomes closer to 1 as a participant's own rating of the PROP solution increases. If we decide through precautionary preferences, we should expect that a t-test would reveal that nontasters are equally if not more harm averse than tasters, and nontasters should reveal no such linear relationship. Second, this work will allow us to compare whether people are able to accurately predict their choices in this task: is there a disconnect between anticipated behavior and actual behavior in this task? Finally, this study will provide estimates of effect size to inform power analyses for the actual task, which I describe now.

Studies 2a and 2b: deciding under phenomenological uncertainty

As described above, this study uses an adapted version of a task that models harm aversion for self and other (Crockett et al., 2014). In the original task, participants make decisions between different amounts of money or electric shocks for themselves or another person. Using a thresholding procedure, the subjective intensity of each shock was matched for both participants. To answer the research questions detailed above, however, Study 2a uses a stimulus that is subjectively very different to both participants—PROP solution.

Before beginning the harm aversion task, participants will determine their taster status and get oriented to the scale using imagined and remembered sensations, as described above. Following a similar procedure as Crockett et al. (2014), both participants will be introduced to the harm aversion task and assigned to either the

Decider or Receiver role. In this case, both participants will be assigned to be the Decider. Since the Receiver must always be a taster, this allows us to avoid pretesting and recruiting a special population, while still maintaining the belief for both participants that the other is the Receiver.

Using a procedure described in detail elsewhere (Crockett et al., 2014, 2017), we will generate 108 trials to efficiently estimate the harm aversion parameters described above. Before beginning the harm aversion task, participants will be reminded of the rating they gave for the PROP solution before learning the apparent rating that the Receiver gave for the PROP solution (an 81 on the gLMS, based on pretest data described below). Participants will complete the task, receive they payout from a randomly selected trial, and leave separately to maintain anonymity.

We will analyze the data in the same way for the real and hypothetical task, and either pattern of results would prove interesting and suggest a number of possible follow-up studies. If nontasters represent the value of an option through empathic projection or simulation, such that they are less harm averse than tasters, future work might explore how we might minimize this egocentric bias in nontasters.

On the other hand, if nontasters are more harm averse than tasters, this would suggest that they are not necessarily representing the value of the outcome for tasters, but rather deciding through precautionary preferences. Future work might explore how nontasters may estimate the value of an outcome in absence of direct experience. In study 3, we explore how different kinds of testimony might provide such information.

Importantly, participants may respond differently in the real and hypothetical tasks. Since there is no financial cost to harm aversion in the hypothetical task,

participants in this study may be more precautionary. This would be in line with past work showing that participants making real decisions are more willing to inflict harm for profit compared to those making hypothetical decisions (FeldmanHall et al., 2012). Thus, we might expect nontasters to respond according to precautionary preferences in the hypothetical task and empathic projection in the real task.

At this point, it is worth considering a potential confound in the task as described: tasters and nontasters differ in more than just their sensitivity to PROP. In fact, taster status predicts some surprising social behavior. For example, supertasters are more reactive than nontasters to inductions of anger (Macht & Mueller, 2007) and show increased reactivity, as measured by startle eye-blink response (Herbert et al., 2014). Thus, it may be that any apparent difference in harm aversion across taster status may not reflect decision-making under phenomenological uncertainty *per se*, but rather differences between tasters and nontasters more broadly.

To account for this, Study 2b aims to replicate Study 2a, but we will randomly assign participants to be "tasters" or "nontasters." Rather than PROP solution, participants will sample either a concentrated or dilute solution of quinine, which we will present to the participants as PROP. Participants who we assign to be "tasters" will sample a more concentrated solution pretested to be at a similar level as taster ratings of PROP, while participants who we assign to be "nontasters" will sample a dilute solution. Thus, the experience of participants in Study 2b will mirror the experience of tasters and nontasters in Study 2a, but this will be through random assignment rather than natural genetic variation. The procedures and anticipated results for Studies 2a and 2b are otherwise identical.

If the pattern of results is the same whether we explore phenomenological uncertainty through PROP sensitivity or through random assignment and deception using quinine, then we will continue future work in this area using PROP. If the results of studies 2a and 2b do differ, however, such that tasters and nontasters systematically differ in ways independent of our task, we will continue future work on the topic using quinine as apparent PROP solution.

Study 3 – The Role of Testimony in Resolving Phenomenological Uncertainty

Given that we can never know what an experience is like for another person, we have limited options for representing the value of that experience. Though we might simulate what that experience would be like for ourselves, using that value as a proxy, we also have access to information about that experience in the form of testimony. Put another way, other people can tell us what an experience is like for them. In my third study, I explore whether some kinds of testimony might be more effective than others.

This work can also help weigh in on a theoretical question of interest: to what extent are phenomenological uncertainty and impact uncertainty distinct concepts? As described above, impact uncertainty is uncertainty surrounding how an outcome would affect the wellbeing of another person, while phenomenological uncertainty is uncertainty surrounding what it is like to actually experience that outcome. In our theoretical model (Chituc, Paul, & Crockett, In Prep), we propose that resolving phenomenological uncertainty also resolves impact uncertainty, but not vice versa. Put another way, knowing what it is subjectively like to experience an outcome provides information about

how good or bad that outcome is, but knowing how good or bad an outcome is does not necessarily give information about what that outcome is subjectively like.

To test this, we will recruit a sample of tasters, explain to them the difference between tasters and nontasters, and ask them to sample and rate the most concentrated solution of PROP used above. Next, we will ask participants to provide two different kinds of testimony about this experience: *phenomenological testimony* and *impact testimony*. To elicit phenomenological testimony, we will ask participants the following. "Imagine that you are talking with a nontaster who tried the same solution and found it completely neutral. Without describing how pleasant or unpleasant it was, how would you describe your experience so that the other person understood what it was like for you?" To elicit impact testimony, we will ask participants the following. "Imagine that you are talking to someone who has been unable to taste or smell anything since birth. How would you describe your experience so that the other person understood how pleasant or unpleasant it was for you?"

Next, we will run an online study on a separate sample of participants based on this first part. Each participant will start by familiarizing themselves with the gLMS using the imagined and remembered sensations as described above. Next, they will read that someone had tasted a solution, rated it using the scale they just learned, and described their experience. Participants will read either the phenomenological testimony or impact testimony given by one participant from the first part of the study, and their task will be to predict the rating that participant gave using the gLMS. Finally, we will ask participants to rate to agree or disagree with the following statements. To assess to what extent phenomenological uncertainty had been resolved, we will ask: "I am confident that

I know what it was like for the other person to taste this solution." To assess to what extent impact uncertainty had been resolved, we will ask: "I am confident that I know how pleasant or unpleasant it was for the other person to taste this solution."

To analyze this data, we will first compute a difference score, which will be the absolute value of the difference between the rating the first participant gave to the solution and the predicted rating that the second participant gave based on the first participant's testimony. We predict that phenomenological testimony will be just as effective as impact testimony, if not better, at allowing participants to predict ratings.

To test our second prediction, that resolving phenomenological uncertainty also resolves impact uncertainty, but not vice versa, we predict the following pattern of results. When rating to what extent impact uncertainty had been resolved, there should be no significant difference between participants in the phenomenological testimony and impact testimony conditions. There should be a significant difference, however, when rating to what extent phenomenological uncertainty had been resolved, such that participants in the impact testimony condition will be significantly less confident that they know what it was like for the first participant to taste the solution.

These findings may open a number of avenues for future work. In either of the predicted pattern of results for Studies 1 – 2b, testimony might be used as an intervention in future harm aversion tasks. If participants decide through empathic projection and simulation, such that nontasters are less harm averse than tasters, we might expect that providing nontasters with phenomenological testimony might lead them to be more harm averse, having κ_{other} more closely aligned with those of tasters. If nontasters decide through precautionary preferences, expressing more harm aversion than tasters, we might

similarly expect that phenomenological testimony will produce κ_{other} more closely aligned with those of tasters, as well.

Finally, we might be able to model a nontasters ability to represent the value of an outcome to nontasters using a learning task, following past work that used the harm aversion task as a way of studying moral inferences (Siegel et al., 2018). Instead of making a series of decisions between different amounts of money and different quantities of PROP solution, participants in a study like this might predict the choices a taster makes for themselves ("Self" trials in the task described above). Here, we might test how quickly tasters and nontasters learn the preferences of a taster in this task, and whether testimony might lead nontasters to learn these preferences more quickly and accurately.

Study 4

In a final line of work, we aim to explore two questions. First, what inferences do people make about the moral character of someone who is deciding under phenomenological uncertainty? And second, to what extent does this differ from inferences we make about someone deciding under other kinds of ignorance or uncertainty?

There is a substantial body of work which explores how ignorance is morally exculpatory, at least in the case of harms (see e.g.: Cushman, 2008; Kissinger-Knox et al., 2018; Young & Saxe, 2011). A nontaster in our harm aversion task, however, does not so cleanly map on to cases of ignorance explored in the literature. There is no outcome uncertainty, since participants know that the solution contains PROP. Furthermore, it's

not clear that there is impact uncertainty, since participants know how the receiver rated the PROP solution on the gLMS.

Thus, this framework allows us to test, in a different way, to what extent impact and phenomenological uncertainty are distinct. Since Deciders know how Receivers rated PROP on the gLMS, they are not in the same position of ignorance someone might be if they were unaware that they were serving a dish with peanuts to someone who had an allergy (c.f. Young & Saxe, 2011). If there is no distinction between impact and phenomenological uncertainty, or if the subjective character of an experience is not a morally relevant consideration, then it should be enough to know that an outcome causes harm. From this, we should expect that nontasters who make selfish decisions in our task would be judged just as harshly as tasters who make selfish decisions, since there the nontaster has no ignorance that might be exculpatory. If, however, these two kinds of uncertainty are distinct, we should predict that nontasters would be judged less harshly than tasters when making selfish decisions.

To test this basic idea, I propose a vignette study that can be conducted online. This study has three conditions: impact uncertainty, phenomenological uncertainty, and no uncertainty. As an example of phenomenological uncertainty, we are use experiencing the loss of a close loved one (e.g. Ruttan et al., 2015). All three conditions start and end in the same way, and the middle differs as marked.

> Imagine that a student is struggling to complete a major assignment after the sudden loss of their mother. After class one day, the student approaches their professor to ask for an extension.
>
> **Impact Uncertainty:** The student does not mention that their mother had recently passed, but they say that they've been struggling to finish the assignment. The professor is sympathetic to the student, even though they aren't sure why the student is struggling.

> **Phenomenological Uncertainty:** The student mentions that their mother had recently passed, and they say that they've been struggling to finish the assignment. The professor is sympathetic to the student, even though they have not personally experienced the unexpected passing of a close love one.
> **No Uncertainty:** The student mentions that their mother had recently passed, and they say that they've been struggling to finish the assignment. The professor is sympathetic to the student, since they had experienced the unexpected passing of a close loved one years earlier.
>
> While the professor believes that the student is genuinely struggling, they ultimately decide against granting an extension, and the student does poorly in the class.

After reading one of these vignettes, participants will rate the moral character of the professor on a scale ranging from "very immoral" (-50) to "very moral" (+50). We will analyze this data using a one-way ANOVA, with pairwise comparisons analyzed using an independent t-test.

In line with past work, we predict that participants will judge the professor least harshly in the impact uncertainty condition and most harshly in the no uncertainty condition. If phenomenological uncertainty is distinct from impact uncertainty, however, we should predict that participants in this condition should judge the professor less harshly than in the no uncertainty condition. If phenomenological uncertainty adds no morally relevant ignorance, however, then we should expect no difference between this condition and the no uncertainty condition.

Though this simple vignette study is just a first step, follow-ups might test this question more rigorously using the task described above. In a 3 (Uncertainty: Impact, Phenomenological, and None) x 2 (Harm aversion: high vs. low) design, we might ask participants to rate the moral character of a Decider in this task. The uncertainty condition follows the same logic as described above. In the impact uncertainty condition, participants will read that Deciders did not know the taster status of the Receiver. In the

phenomenological uncertainty condition, participants will read that Decider knows the status of the Receiver as well as the rating the Receiver provided, but the Decider is a nontaster and experiences PROP as neutral. In the no uncertainty condition, participants will read that the Decider knows the status of the Receiver as well as the rating they provided, but is a taster and experiences PROP as intensely bitter. Thus, we might see how these different kinds of uncertainty influence the moral inferences we make about someone who requires either the minimum or maximum amount of money (10 cents vs. $20) to administer 20 doses of PROP. Such a design, though less naturalistic, has a few benefits. First, it more cleanly ensures that impact uncertainty is matched across condition, since both tasters and nontasters know how the receiver rated PROP on the gLMS. Second, it allows us to explore whether there is an asymmetry between praise and blame.

We expect the same basic pattern of results as described above: if phenomenological uncertainty and impact uncertainty are distinct, participants should infer that less harm averse tasters are worse than less harm averse nontasters who make the same series of decisions. If these kinds of uncertainty are not distinct, however, we should expect that participants rate tasters and nontasters with equal harshness.

Proposed Timeline and Conclusion

This prospectus details a number of potential avenues to test basic questions surrounding how we make decisions about experiences we can't imagine or otherwise represent. In doing so, this work would contribute to work in judgment and decision-making, helping to answer how we assign value to an outcome we in principle cannot know? Even more,

this work may help point to interventions we might apply to make these decisions better. For example, if we are subject to an egocentric bias, can phenomenological testimony lead nontasters to decide in similar ways to tasters? This work also contributes to research on decision-making under uncertainty. Phenomenological uncertainty has yet to be empirically explored, and this work will answer whether and to what it extent it is unique from other kinds of uncertainty. And finally, it will contribute to work in moral psychology. How are we judged for failing to understand what it's like to have a different experience of the world? Is this required for moral responsibility?

In large part, the studies I propose here, and the branches I propose for future work, serves as a contingency plan given the ongoing COVID-19 pandemic and the uncertainty surrounding my ongoing ability to conduct human subjects research in the lab. While I am able to start or resume data collection for projects that must be conducted in-person in the coming spring, as of now there is no clear sense on when it may be safe to do so.

In the interim, the simpler designs, like Study 4, can be conducted online. Should initial results be promising, and should there be a large delay in the ability to collect in-person data, it is possible to adapt the moral inference task sketched above using an online sample. We might also adapt Study 3 to be conducted online, perhaps by recruiting participants who have eaten durian fruit (see: Paul, 2014, p15). Thus, even if there is a large delay in my ability to collect data in the lab, I have sketched a number of viable lines of research that I can pursue. While it may not be possible to include all of this work in my dissertation, the studies proposed here nonetheless provide enough opportunity to sufficiently explore the topic of phenomenological uncertainty in my dissertation work.

References

Adler, M. D. (2014). Extended Preferences and Interpersonal Comparisons: A New Account. *Economics and Philosophy*, *30*(02), 123–162. https://doi.org/10.1017/S0266267114000133

Barron, H. C., Dolan, R. J., & Behrens, T. E. J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience*, *16*(10), 1492–1498. https://doi.org/10.1038/nn.3515

Bartoshuk, Linda M., Duffy, V. B., & Miller, I. J. (1994). PTC/PROP tasting: Anatomy, psychophysics, and sex effects. *Physiology & Behavior*, *56*(6), 1165–1171. https://doi.org/10.1016/0031-9384(94)90361-1

Bartoshuk, Linda M., & Snyder, D. J. (2004). Psychophysical Measurement of Human Taste Experience. In E. M. Stricker & S. C. Woods (Eds.), *Neurobiology of Food and Fluid Intake* (Vol. 14, pp. 89–107). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48643-1_5

Bartoshuk, L.M, Duffy, V. B., Green, B. G., Hoffman, H. J., Ko, C.-W., Lucchina, L. A., Marks, L. E., Snyder, D. J., & Weiffenbach, J. M. (2004). Valid across-group comparisons with labeled scales: The gLMS versus magnitude matching. *Physiology & Behavior*, *82*(1), 109–114. https://doi.org/10.1016/j.physbeh.2004.02.033

Birnbaum, M. H. (1999). How to show that 9 > 221: Collect judgments in a between-subjects design. *Psychological Methods*, *4*(3), 243.

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, *8*(1), 15958. https://doi.org/10.1038/ncomms15958

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*(48), 17320–17325. https://doi.org/10.1073/pnas.1408988111

Crockett, M. J., Siegel, J. Z., Kurth-Nelson, Z., Dayan, P., & Dolan, R. J. (2017). Moral transgressions corrupt neural representations of value. *Nature Neuroscience*, *20*(6), 879–885. https://doi.org/10.1038/nn.4557

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. https://doi.org/10.1016/j.cognition.2008.03.006

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, *33*(1), 67–80. https://doi.org/10.1007/s00199-006-0153-z

Ditschun, T. L., & Guinard, J.-X. (2004). Comparison of New and Existing Methods for the Classification of Individuals According to 6-N Propylthiouracil (prop) Taster Status. *Journal of Sensory Studies*, *19*(2), 149–170. https://doi.org/10.1111/j.1745-459X.2004.tb00142.x

Drewnowski, A., Henderson, S., Shore, A., & Barett-Fornell, A. (1997). Nontasters, Tasters, and Supertasters of 6-n-Propylthiouracil (PROP) and Hedonic Response

to Sweet. *Physiology & Behavior*, *62*(3), 649–655. https://doi.org/10.1016/S0031-9384(97)00193-5

Duffy, V., Peterson, J., & Bartoshuk, L. (2004). Associations between taste genetics, oral sensation and alcohol intake. *Physiology & Behavior*, *82*(2–3), 435–445. https://doi.org/10.1016/j.physbeh.2004.04.060

FeldmanHall, O., Mobbs, D., Evans, D., Hiscox, L., Navrady, L., & Dalgleish, T. (2012). What we say and what we do: The relationship between real and hypothetical moral choices. *Cognition*, *123*(3), 434–441. https://doi.org/10.1016/j.cognition.2012.02.001

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Goldman, A. I. (1995). Simulation and Interpersonal Utility. *Ethics*, *105*(4), 709–726. https://doi.org/10.1086/293749

Hammond, P. J. (1977). DUAL INTERPERSONAL COMPARISONS OF UTILITY AND THE WELFARE ECONOMICS OF INCOME DISTRIBUTION. *Journal of Public Economics*, *7*(1), 51–71.

Harsanyi, J. C. (1977). Morality and the Theory of Rational Behavior. *Social Research*, 623–656.

Hausman, D. M. (1995). The Impossibility of Interpersonal Utility Comparisons. *Mind*, *104*(415), 473–490. https://doi.org/10.1093/mind/104.415.473

Hayes, J. E., Allen, A. L., & Bennett, S. M. (2013). Direct comparison of the generalized visual analog scale (gVAS) and general labeled magnitude scale (gLMS). *Food*

Quality and Preference, *28*(1), 36–44. https://doi.org/10.1016/j.foodqual.2012.07.012

Herbert, C., Platte, P., Wiemer, J., Macht, M., & Blumenthal, T. D. (2014). Supertaster, super reactive: Oral sensitivity for bitter taste modulates emotional approach and avoidance behavior in the affective startle paradigm. *Physiology & Behavior*, *135*, 198–207. https://doi.org/10.1016/j.physbeh.2014.06.002

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic Social Cognition: Empathic Accuracy in Mixed-Sex Dyads. *Journal of Personality and Social Psychology*, *59*(4), 730–742.

Kappes, A., Nussberger, A.-M., Faber, N. S., Kahane, G., Savulescu, J., & Crockett, M. J. (2018). Uncertainty about the impact of social decisions increases prosocial behaviour. *Nature Human Behaviour*, *2*(8), 573–580. https://doi.org/10.1038/s41562-018-0372-x

Kappes, A., Nussberger, A.-M., Siegel, J. Z., Rutledge, R. B., & Crockett, M. J. (2019). Social uncertainty is heterogeneous and sometimes valuable. *Nature Human Behaviour*, *3*(8), 764–764. https://doi.org/10.1038/s41562-019-0662-y

Kissinger-Knox, A., Aragon, P., & Mizrahi, M. (2018). Does Non-Moral Ignorance Exculpate? Situational Awareness and Attributions of Blame and Forgiveness. *Acta Analytica*, *33*(2), 161–179. https://doi.org/10.1007/s12136-017-0339-y

Krueger, J., & Clement, R. W. (1997). Estimates of Social Consensus by Majorities and Minorities: The Case for Social Projection. *Personality and Social Psychology Review*, *1*(4), 299–313. https://doi.org/10.1207/s15327957pspr0104_2

Macht, M., & Mueller, J. (2007). Increased negative emotional responses in PROP supertasters. *Physiology & Behavior*, *90*(2–3), 466–472. https://doi.org/10.1016/j.physbeh.2006.10.011

Marks, L. E., Stevens, J. C., Bartoshuk, L. M., Gent, J. F., Rifkin, B., & Stone, V. K. (1988). Magnitude-matching: The measurement of taste and smell. *Chemical Senses*, *13*(1), 63–87. https://doi.org/10.1093/chemse/13.1.63

Murty, V., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology. General*, *145*(5), 548–558. https://doi.org/10.1037/xge0000158

Paul, L. A. (2014). *Transformative experience* (1st ed). Oxford University Press.

Platt, M. L., & Huettel, S. A. (2008). Risky business: The neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, *11*(4), 398–403. https://doi.org/10.1038/nn2062

Prutkin, J., Duffy, V. B., Etter, L., Fast, K., Gardner, E., Lucchina, L. A., Snyder, D. J., Tie, K., Weiffenbach, J., & Bartoshuk, L. M. (2000). Genetic variation and inferences about perceived taste intensity in mice and men. *Physiology & Behavior*, *69*(1–2), 161–173. https://doi.org/10.1016/S0031-9384(00)00199-2

Robbins, L. (1935). *Essay on the Nature and Significance of Economic Science (1st. Ed. 1932).* Macmillan.

Ruttan, R. L., McDonnell, M.-H., & Nordgren, L. F. (2015). Having "been there" doesn't mean I care: When prior experience reduces compassion for emotional distress. *Journal of Personality and Social Psychology*, *108*(4), 610–622. https://doi.org/10.1037/pspi0000012

Sen, A. (1974). Informational bases of alternative welfare approaches: Aggregation and income distribution. *Journal of Public Economics*, *3*(4), 387–403. https://doi.org/10.1016/0047-2727(74)90006-1

Shanton, K., & Goldman, A. (2010). Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 527–538. https://doi.org/10.1002/wcs.33

Siegel, J. Z., Mathys, C., Rutledge, R. B., & Crockett, M. J. (2018). Beliefs about bad people are volatile. *Nature Human Behaviour*, *2*(10), 750–756. https://doi.org/10.1038/s41562-018-0425-1

Snyder, D. J., Sims, C. A., & Bartoshuk, L. M. (2015). Psychophysical Measures of Human Oral Sensation. In R. L. Doty (Ed.), *Handbook of Olfaction and Gustation* (pp. 749–774). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118971758.ch34

Sunstein, C. R. (2005). *Laws of fear: Beyond the precautionary principle* (Vol. 6). Cambridge University Press.

Suppes, P., & Winet, M. (1955). An Axiomatization of Utility Based on the Notion of Utility Differences. *Management Science*, *1*(3), 259–270.

Tepper, B. J. (1998). 6-n-Propylthiouracil: A Genetic Marker for Taste, with Implications for Food Preference and Dietary Habits. *The American Journal of Human Genetics*, *63*(5), 1271–1276. https://doi.org/10.1086/302124

Tepper, B. J., Christensen, C. M., & Cao, J. (2001). Development of brief methods to classify individuals by PROP taster status. *Physiology & Behavior*, *73*(4), 571–577. https://doi.org/10.1016/S0031-9384(01)00500-5

Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, *120*(2), 202–214. https://doi.org/10.1016/j.cognition.2011.04.005

Moral Life Under Phenomenological Uncertainty

Vladimir Chituc

Introduction (short version)

In my dissertation work, I aim to explore a commonplace and underappreciated aspect of decision-making: we must often make decisions about an outcome we haven't experienced and cannot imagine. In my dissertation work, I argue that these decisions involve what we call *phenomenological uncertainty,* in that we are unsure of what it is subjectively like to experience an outcome. Here, we ask how this uncertainty is resolved, how it differs from other kinds of uncertainty, and how we might resolve it better.

Far from being a unified concept, there are a number of varieties of uncertainty (Kappes et al., 2019) that affect our decisions in different ways. Consider two kinds of uncertainty: uncertainty around the likelihood of an outcome, or *outcome uncertainty* (e.g. Platt & Huettel, 2008), and uncertainty about how that outcome might affect other people, or *impact uncertainty* (Kappes et al., 2018). In our theoretical model, we argue that impact and phenomenological uncertainty are distinct. To briefly summarize our view, one way we assign an outcome value is through reference to experience. Thus, knowing what an experience is like allows us to assign it value, but knowing the value of an experience does not tell us what it's like. If phenomenological uncertainty is a distinct kind of uncertainty, how, then, do we make decisions that involve it?

One possibility is through *empathic projection* or simulation (Shanton & Goldman, 2010). Briefly, we use our own experience as a benchmark when considering the experiences of others. In contrast, we may decide through *precautionary preferences,* making decisions as if the worst outcome were true given certain kinds of uncertainty (Crockett et al., 2014; Kappes et al., 2018).

Overview of Proposed Studies

To study these topics in the lab, we make use of a phenomenon identified in the psychophysics literature—some people ("tasters") are able to taste an extremely bitter compound, 6-n-propylthiouracil (PROP), while others ("nontasters") are not. From this, I propose a number of studies. First, I describe a hypothetical (Study 1) and real (Studies 2a & 2b) harm aversion task. Next, I describe work exploring the role of testimony (Study 3), which will help explore to what extent impact and phenomenological uncertainty are distinct. Finally, I propose a vignette study to explore how we make moral judgments in cases of phenomenological uncertainty (Study 4).

Studies 1, 2a, and 2b: deciding under phenomenological uncertainty

Studies 1, 2a, & 2b adapts a harm aversion task (Crockett et al., 2014) with two roles: a "Decider" and a "Receiver." The Decider will make a series of choices between higher quantities of PROP solution for more money and lower quantities of PROP solution for less money. While the Decider always receives the money, the Decider must sometimes choose between options where they sample PROP or when the Receiver samples PROP. From the choices in this task, we can calculate a harm aversion parameter (κ) between 0 and 1, where higher values mean greater harm aversion.

As described above, while the Decider in this task will sometimes be a taster and other times not, the apparent Receiver in this task will always be a taster. This will allow us to compare κ for both tasters and nontasters, which will allow us to see how harm aversion varies based on whether phenomenological uncertainty is present (when nontasters decide for tasters) or absent (when tasters decide for tasters). We will classify participants according to taster status using an established method (Tepper et al., 2001)

which is based on how they rate the PROP solution using the generalized labeled magnitude scale (gLMS; Bartoshuk et al., 2004), which ranges from "no sensation" (0) to "strongest imaginable sensation of any kind" (100).

In our hypothetical task (Study 1), we will recruit participants and present them with a hypothetical version of our main task as described above. After sampling and rating the PROP solution, participants will make 20 decisions between different amounts of money and different numbers of doses of PROP solution to be delivered to a hypothetical Receiver. The second set of studies (2a & 2b) use an adapted version of a task that models harm aversion for self and other (Crockett et al., 2014). Before beginning the real task, participants will determine their taster status and be assigned to either the Decider or Receiver role.

We will analyze the data in the same way for the real and hypothetical task. If we decide through empathic projection, we should expect that tasters are more harm averse than nontasters, and that there is a linear relationship such that κ becomes closer to 1 as a participant's own rating of the PROP solution increases. If we decide through precautionary preferences, we should expect that nontasters are equally if not more harm averse than tasters, and nontasters should reveal no such linear relationship.

There is, however, a potential confound in the task as described: taster status predicts some surprising social behavior (e.g. Macht & Mueller, 2007; Herbert et al., 2014). Thus, it may be that any apparent difference in harm aversion across taster status may not reflect decision-making under phenomenological uncertainty. To account for this, Study 2b aims to replicate Study 2a, but we will randomly assign participants to be "tasters" or "nontasters." Rather than PROP solution, participants will sample solutions

of quinine that are either concentrated ("taster") or dilute ("nontaster"). The procedures and anticipated results for Studies 2a and 2b are otherwise identical.

Studies 3 and 4 - Testimony and Moral Judgments

While we do not have direct access to the experiences of others, we do learn information through testimony. Put another way, other people can tell us what an experience is like for them. In Study 3, I explore whether some kinds of testimony might be more effective than others, and in doing so, I ask whether phenomenological uncertainty and impact uncertainty are distinct concepts. In our theoretical model, resolving phenomenological uncertainty resolves impact uncertainty, but not vice versa.

To test this, we will recruit a sample of tasters and ask them to sample and rate the most concentrated solution of PROP used above. Next, we will ask participants to provide two different kinds of testimony about this experience: either about what an experience is *like* without reference to how positive or negative it is ("phenomenological testimony*"*), or testimony about how positive or negative it is without reference to what it is like ("impact testimony"). A separate sample of participants will predict how participants rated the PROP solution based on this testimony. Then, we will assess to what extent phenomenological uncertainty and impact uncertainty has been resolved.

We predict that phenomenological testimony will be just as effective as impact testimony, if not better, at allowing participants to predict ratings. We also predict that phenomenological testimony will resolve both impact and phenomenological uncertainty, but impact testimony will only resolve impact uncertainty.

In Study 4, we aim to explore two questions. First, what inferences do people make about the moral character of someone who is deciding under phenomenological

uncertainty? And second, to what extent does this differ from inferences we make about someone deciding under other kinds of ignorance or uncertainty?

There is a substantial body of work which explores how ignorance is morally exculpatory, at least in the case of harms. A nontaster in our harm aversion task, however, does not so cleanly map on to cases of ignorance explored in the literature, since participants know how the receiver rated the PROP solution on the gLMS.

This framework allows us to test, in a different way, whether impact and phenomenological uncertainty are distinct. If there is no difference between the two, or if the subjective character of an experience is not a morally relevant consideration, then it should be enough to know that an outcome causes harm. To test this basic idea, I propose a vignette study that can be conducted online. As an example of phenomenological uncertainty, we are use experiencing the loss of a close loved one (e.g. Ruttan et al., 2015). Participants will read one of three conditions which describe a professor who denies an extension to a student who suddenly lost a parent. In some cases, the professor will not know why the student is struggling (impact uncertainty). In others, the professor will know, but they either haven't experienced the sudden loss of a parent themselves (phenomenological uncertainty), or they have (no uncertainty).

After reading one of these vignettes, participants will rate the moral character of the professor, and we will ask participants to rate how confident they are in their rating. In line with past work, we predict that participants will judge the professor least harshly in the impact uncertainty condition and most harshly in the no uncertainty condition. If phenomenological uncertainty is distinct from impact uncertainty, however, we should predict that participants in this condition should judge the professor less harshly than in

the no uncertainty condition. If phenomenological uncertainty adds no morally relevant ignorance, however, then we should expect no difference between this condition and the no uncertainty condition.

Proposed Timeline and Conclusion

This prospectus details a number of potential avenues to test basic questions surrounding how we make decisions about experiences we can't imagine or otherwise represent. In doing so, this work would contribute to work in judgment and decision-making, helping to answer how we assign value to an outcome we in principle cannot know? Even more, this work may help point to interventions we might apply to make these decisions better.

I aim to begin collecting lab data (Studies 1 – 2b) in the fall or spring, but this may be substantially delayed given the ongoing COVID-19 pandemic. Given this, I am able to start online data collection this summer (Study 4). Should initial results be promising, and should there be a large delay in the ability to collect in-person data, we might extend this work in follow up studies while adapting Study 3 for online use. Should this be the case, we will begin collecting in person lab data as soon as it is safe and feasible to do so. Thus, even if there is a large delay in my ability to collect data in the lab, I have sketched a number of viable lines of research that I can pursue.

References

Bartoshuk, L. M., Duffy, V. B., Green, B. G., Hoffman, H. J., Ko, C.-W., Lucchina, L. A., Marks, L. E., Snyder, D. J., & Weiffenbach, J. M. (2004). Valid across-group comparisons with labeled scales: The gLMS versus magnitude matching. *Physiology & Behavior*, *82*(1), 109–114. https://doi.org/10.1016/j.physbeh.2004.02.033

Crockett, M. J., Kurth-Nelson, Z., Siegel, J. Z., Dayan, P., & Dolan, R. J. (2014). Harm to others outweighs harm to self in moral decision making. *Proceedings of the National Academy of Sciences*, *111*(48), 17320–17325. https://doi.org/10.1073/pnas.1408988111

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353–380. https://doi.org/10.1016/j.cognition.2008.03.006

Herbert, C., Platte, P., Wiemer, J., Macht, M., & Blumenthal, T. D. (2014). Supertaster, super reactive: Oral sensitivity for bitter taste modulates emotional approach and avoidance behavior in the affective startle paradigm. *Physiology & Behavior*, *135*, 198–207. https://doi.org/10.1016/j.physbeh.2014.06.002

Kappes, A., Nussberger, A.-M., Faber, N. S., Kahane, G., Savulescu, J., & Crockett, M. J. (2018). Uncertainty about the impact of social decisions increases prosocial behaviour. *Nature Human Behaviour*, *2*(8), 573–580. https://doi.org/10.1038/s41562-018-0372-x

Kappes, A., Nussberger, A.-M., Siegel, J. Z., Rutledge, R. B., & Crockett, M. J. (2019). Social uncertainty is heterogeneous and sometimes valuable. *Nature Human Behaviour*, *3*(8), 764–764. https://doi.org/10.1038/s41562-019-0662-y

Kissinger-Knox, A., Aragon, P., & Mizrahi, M. (2018). Does Non-Moral Ignorance Exculpate? Situational Awareness and Attributions of Blame and Forgiveness. *Acta Analytica*, *33*(2), 161–179. https://doi.org/10.1007/s12136-017-0339-y

Macht, M., & Mueller, J. (2007). Increased negative emotional responses in PROP supertasters. *Physiology & Behavior*, *90*(2–3), 466–472. https://doi.org/10.1016/j.physbeh.2006.10.011

Platt, M. L., & Huettel, S. A. (2008). Risky business: The neuroeconomics of decision making under uncertainty. *Nature Neuroscience*, *11*(4), 398–403. https://doi.org/10.1038/nn2062

Ruttan, R. L., McDonnell, M.-H., & Nordgren, L. F. (2015). Having "been there" doesn't mean I care: When prior experience reduces compassion for emotional distress. *Journal of Personality and Social Psychology*, *108*(4), 610–622. https://doi.org/10.1037/pspi0000012

Shanton, K., & Goldman, A. (2010). Simulation theory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(4), 527–538. https://doi.org/10.1002/wcs.33

Tepper, B. J., Christensen, C. M., & Cao, J. (2001). Development of brief methods to classify individuals by PROP taster status. *Physiology & Behavior*, *73*(4), 571–577. https://doi.org/10.1016/S0031-9384(01)00500-5

Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, *120*(2), 202–214. https://doi.org/10.1016/j.cognition.2011.04.005