

## 1 Introduction

Compared to other aspects of Kant's practical philosophy, Kant's theory of conscience remains relatively unexplored in the secondary literature on his work. This is no doubt due, at least in part, to the fact that in the *Groundwork of the Metaphysics of Morals* (henceforth *Groundwork*) and the *Critique of Practical Reason*, Kant's two most widely read works on ethics, conscience plays very little role.<sup>1</sup>

However, Kant has extended discussions of conscience in three of his lesser-read works: "On the Miscarriage of all Philosophical Attempts in Theodicy" (henceforth "Miscarriage"), *Religion within the Boundaries of Mere Reason* (henceforth *Religion*), and the *Metaphysics of Morals*. There are also many unpublished notes in which Kant discusses conscience, and it may be conjectured, on the basis of extant copies of students' notes, that conscience was a frequent topic in Kant's lectures.

As commentators have begun to pay closer attention to these lesser-read works, the literature on Kant's theory of conscience has begun to develop. It is thanks to this development that this Element is possible: there is an emerging consensus that Kant's theory of conscience is important both in its own right and insofar as it can help to correct various misunderstandings that have become part of the standard view of Kant's ethical thought.

The main body of this Element is divided into two sections. Section 2 focuses on exegesis of Kant's ethics. One of the overarching theses of this section of the Element is that, although many of Kant's claims about conscience are *prima facie* inconsistent, a close examination of context generally can dissolve apparent contradictions. Section 3 of the Element focuses on philosophical issues in Kantian ethics. One of the overarching theses of this section is that many positions traditionally associated with Kantian ethics, including the denial of moral luck, the nonaccidental rightness condition, and the guise of the objectively good, are at variance with Kant's ethics.

I would like to conclude this section with a personal note about why I thought it appropriate to dedicate this particular work to my teachers. I do not know whether this is a common experience, but in my case, my teachers have believed in me, supported me, and spurred me on to do better work notwithstanding having what, on reflection, I consider a poor evidentiary base. Indeed, that even occurred in this Element, the first draft of which was garbage: I rewrote it only because one of my

---

<sup>1</sup> I am able to find only three mentions of conscience in these two works, each time merely in passing: GMS, AA 04: 404.20; GMS, AA 04: 422.19; and KpV, AA 05: 098.14. Note that all references to Kant's work use the standard academy pagination. The academy pagination can usually be found in the margins of translations of these texts. I have consulted, when possible, the Guyer/Wood Cambridge Edition translations. But all translations (and errors) here are my own.

teachers believed in me enough to think I could do better. I leave it to them to determine whether their continued support of me could be in accordance with the strict standards of belief associated with a Kantian faculty of conscience. But I do hope to convey some small measure of my heartfelt gratitude.

## 2 Kant's Ethics

As noted in Section 1, this part of the Element is exegetical. I have divided what I take to be Kant's most important remarks on conscience into five themed subsections. In subsection 2.1, I examine Kant's definitions of conscience and the functions he attributes to it. In subsection 2.2, I examine his position on errors of conscience. In subsection 2.3, I examine Kant's ideas about acting in accordance with conscience. In subsection 2.4, I examine the ways in which Kant uses ideas about conscience to illustrate his commitment to strict evidentialism and freedom of conscience. In subsection 2.5, I examine how Kant's ideas about conscience bear on his theism.

### 2.1 Definitions

Conscience is, for Kant, a mental faculty, or at the very least a capacity that arises as the result of the interplay of other mental faculties. As such, I begin by delineating the functions Kant assigns to it. The most essential function of conscience, for Kant, is to determine whether an agent has acted in accordance with the moral law. In the *Metaphysics of Morals*, Kant compares conscience to a criminal court. Kant describes conscience as having a second-order function, one that subsumes the activities of other mental faculties. According to Kant, the faculty of moral understanding determines the moral rule, playing the role of moral legislator; the faculty of judgment decides whether a given action, as falling under this law, has taken place and then imputes the action to the agent; and the faculty of reason then either acquits or condemns the agent.<sup>2</sup> Kant says that "all of this happens before a court . . . called a court of justice" and uses this to define conscience: "The consciousness of an *inner court of justice* in the human ('before which his thoughts themselves accuse or excuse one another') is **conscience**."<sup>3</sup> From this it may be seen that, for Kant, whereas a criminal court determines whether a criminal is innocent or guilty of a given crime, the court of conscience determines whether an agent has acted in conformity with or contrary to the moral law in a given instance.

Kant assigns a similar function to conscience in the *Religion*. In the *Religion*, Kant says that "[t]his spirit [viz., the Holy Ghost, as opposed to God the father

<sup>2</sup> MS, AA 06: 437.32–438.08.    <sup>3</sup> MS, AA 06: 438.08–12.

or God the son] . . . is at the same time the actual judge of men (through their conscience).”<sup>4</sup> He then distinguishes between two kinds of judging: (1) judging in regard to merit and lack of merit, and (2) judging in regard to guilt and innocence. According to Kant, God “in his son” passes judgment of the first kind, and this kind of judgment involves interpersonal comparison. The second kind of judgment, by way of contrast, is what God as the Holy Spirit is responsible for, and such a judgment is “about *one and the same* person before a court (conscience).”<sup>5</sup>

Kant’s parsing of the trinity is unimportant for current purposes. What is important is that conscience, for Kant, is involved in judging whether an agent is worthy of reproach for having infringed on duty. To be clear, this is not to be understood as reproach in the eyes of others. Kant’s idea is not that conscience judges whether others’ censure is justified. Rather, conscience judges whether an agent’s behavior has measured up to the strict standard of the moral law. Others might engage in blame behavior or resent someone for infringing on a duty. But blame, for Kant, is fundamentally personal; there is, for Kant, a fact of the matter, one independent of any blame-practices or interpersonal feelings or judgments, about whether an agent is blameworthy, and it is this fact that the faculty of conscience uncovers. (There also might be a fact about whether an agent deserves merit. But that is not, according to this passage, the domain of conscience.)

There are many unpublished notes in which Kant assigns a similar function to conscience.<sup>6</sup> For example, in a note dated to 1783–4, Kant gives a twofold definition of conscience.<sup>7</sup> He says that conscience is “the capacity to become conscious of the rightfulness or wrongfulness of one’s own actions” and also “the inner standing of the evaluatory capacity, as that of a judge, to draw us to account regarding the permission of our actions.”<sup>8</sup> In a note dated to 1773–8,

<sup>4</sup> RGV, AA 06: 145.29–34.    <sup>5</sup> RGV, AA 06: 146.19.

<sup>6</sup> These notes are taken from a variety of sources. They are taken from textbooks that Kant used for teaching and then scribbled marginal notes in; from copies of Kant’s own works that he went through and annotated; and from loose sheets of paper, old letters and the like, on which Kant wrote down ideas and keywords. They include rough drafts of works intended for publication (including drafts of works that eventually were published, like the *Metaphysics of Morals*) in various degrees of completion; and they include the parts of the project Kant was working on during the last years of his life, published after his death sometimes as a standalone work called the *Opus Postumum*.

<sup>7</sup> The dating of these notes is both imprecise and of questionable accuracy. The standard dating used today (which will be used here) was set out by the Kant scholar Erich Adickes in the first half of the twentieth century. He relied on contextual clues (such as dates on letters for notes written on correspondence); changes in handwriting (both style and ink); placement of text on a page (e.g., if a note is written between the lines of another one, then it must have come later); and other clues to piece together a rough chronological map. The mapping seems to be reasonable, all things considered, but (to use an idea of Kant’s own) it is doubtful whether many would wager much on it.

<sup>8</sup> Refl, AA 18: 579.04–07.

Kant writes that conscience is “the consciousness of the duty to be honest in the imputation of one’s own deed,”<sup>9</sup> and he then goes on, as in the *Metaphysics of Morals*, to compare conscience to a court of law, one in which “understanding [is] the lawgiver, the power of judgment [is] the prosecutor and the advocate, but reason is the judge.”<sup>10</sup> And a short note from 1772–8 says merely: “conscience, which reproaches and imputes.”<sup>11</sup>

A similar lesson may also be drawn from the discussions of conscience that can be found throughout the extant sets of students’ notes from Kant’s lectures.<sup>12</sup> For example, in one set of notes, conscience is defined as “an instinct according to which our actions 1. are imputed and 2. applied to the law 3. as well as judged with rightful force. In general it is the drive in our nature to judge ourselves.”<sup>13</sup> And then, about thirty pages later, in the same set of notes, this is repeated: conscience is claimed to be “an instinct to judge oneself morally.”<sup>14</sup> In another set of lecture notes it is claimed that “there is thus an instinct to judge about our actions . . . and this is conscience.”<sup>15</sup> This claim is then echoed about fifty pages later (in the same set of notes): “conscience is an instinct to judge oneself according to moral laws.”<sup>16</sup> And in a third set of lecture notes it is claimed that conscience is “the capacity to impute to oneself one’s own factum through the law.”<sup>17</sup>

<sup>9</sup> Refl, AA 19: 170.23–24.

<sup>10</sup> Refl, AA 19: 170.26–27. This idea is also found in Kant’s correspondence. For example, see his 1792 letter to Maria von Herbert (Br, AA 11: 333.29–33).

<sup>11</sup> Refl, AA 15: 149.09.

<sup>12</sup> There seems to have been a regular trade in these notebooks when Kant was alive, with students buying and selling them as study guides to accompany Kant’s lectures. As such, it is not known either when or by whom the still surviving notebooks were written. It is not known even how many authors contributed to a single notebook. One thing that does seem apparent is that most of these texts are multiply removed from Kant: Kant gave a lecture; someone took notes during the lecture; someone extrapolated those notes into a text in prose form outside of the lecture; someone copied that text, perhaps emending, extrapolating, and interpreting as they went; someone received that text, perhaps emending again; and so on. By means of careful (and painstaking) detective work, scholars sometimes have been able to arrive at more or less reliable conjectures about the rough origins of one or another of these texts. But it is notable that even if there were verbatim transcripts of Kant’s lectures, it still would be unclear to what extent the views expressed should be taken as Kant’s own. For my part, I would be more than a little apprehensive if someone were to use, without warning, a lecture transcript from a day in one of my classes as the basis of a substantive interpretation of my thought.

<sup>13</sup> V-PP/Powalski, AA 27: 162.04–07. <sup>14</sup> V-PP/Powalski, AA 27: 196.16–17.

<sup>15</sup> V-Mo/Collins, AA 27: 297.05–06. <sup>16</sup> V-Mo/Collins, AA 27: 351.22–23.

<sup>17</sup> V-MS/Vigil, AA 27: 575.27–28. Often when it is claimed that conscience is an instinct, it is also claimed that conscience is not a capacity. This distinction is not always explained. But at least once it is said to consist in the fact that an instinct is not under voluntary control: the text argues that, although agents do have the capacity to judge their own actions voluntarily, conscience does so regardless of agents’ wills (V-Mo/Collins, AA 27: 351.22–30).

However, the inconsistency between the passages in which the texts claim that conscience is not a capacity and the ones in which the texts claim that it is a capacity might be merely verbal.

The function of judging whether an agent is guilty of an obligation infringement is linked up to a second function that Kant assigns to conscience: causing pain. As evidence for this, consider the following passage from Kant's *Metaphysics of Morals*:

The pain which a person would feel from the bites of conscience, although its origin is moral, is nevertheless physical according to its effect, just as peevishness, fear, and every other morbid condition.<sup>18</sup>

According to Kant, pain is the effect of the reproaches (bites) of conscience that an agent feels when s/he fails to fulfill prescriptions (or proscriptions) of the moral law.

A few pages later in the *Metaphysics of Morals*, Kant remarks that conscience is one of several “*aesthetic* and preceding but natural predispositions of the mind (*praedispositio*) to be affected by concepts of duty.”<sup>19</sup> Kant's idea in this passage is that without conscience (and these other predispositions), a being could not be moral to begin with. Kant's wording has suggested to some commentators that conscience, in addition to whatever other roles it may play in Kant's moral psychology, is itself a feeling.<sup>20</sup> Indeed, in the Cambridge Edition of the Works of Immanuel Kant, the word I have translated as “*aesthetic*” (*ästhetisch*) and which appears also in the title of the section is translated in both places as “feeling.” In my view, this interpretation is not well supported by the text. Kant does not explicitly say that conscience is a feeling. But he does explicitly say, in the passage just above, that conscience causes feeling. And he also says explicitly in this passage that conscience is “not of empirical origin” and that conscience has the power “to *affect* the moral feeling through its act.”<sup>21</sup> Moreover, in the passage from which the court metaphor that opened this subsection is taken, Kant says conscience is an “intellectual . . . predisposition.”<sup>22</sup> These claims are in tension with the idea that conscience is itself a feeling. Based on this (i.e., the lack of explicit support and the tension it

---

For example, only about forty pages after the claim that conscience is the capacity of imputation in the passage to which this note is appended, conscience is compared to *apperceptio*:

Conscientia taken generally is the consciousness of oneself like *apperceptio*; in specie it involves the consciousness of my will, my disposition, / to do right, or that the action is right, = hence a consciousness of that which for itself is duty.

(V-MS/Vigil, AA 27: 613.38–614.03)

The reason this would render the inconsistency merely verbal is that *apperceptio*, as used here, is not a faculty that is under voluntary control. This suggests that the word “capacity” is not used univocally in these seemingly inconsistent passages.

<sup>18</sup> MS, AA 06: 394.03–05.   <sup>19</sup> MS, AA 06: 399.10–12.

<sup>20</sup> Wood (2008, chapter 10, section 1).

<sup>21</sup> MS, AA 06: 399.14–15 and 400.30, emphasis mine.   <sup>22</sup> MS, AA 06: 438.24–25.

would create with Kant's other claims about conscience), I suggest that the interpretation of conscience as a feeling is mistaken.<sup>23</sup> I suggest, further, that Kant's claims about conscience being an aesthetic predisposition might be related to the claim, seen earlier in the lecture notes, that conscience is an instinct. The key here is that conscience is not entirely under voluntary control, something I shall return to momentarily.

Kant draws attention to an asymmetry in the feelings caused by conscience. Although Kant thinks that conscience can cause an agent pain (by means of reproaches and "bites of conscience"), he does not think that conscience can cause an agent to feel positive joy: "bliss in the comforting encouragement of one's conscience is not *positive* (as joy) but rather only *negative* (becalming after the foregoing trepidation)."<sup>24</sup> This is because conscience can reach one of only two verdicts: guilty or innocent. If the agent is guilty and s/he is not depraved, this will cause him/her pain. But if the agent is innocent, that is no reward; there is no joy that is attendant upon this verdict, at least according to Kant's moral psychology. Thus, this asymmetry in the feelings caused by conscience is related to the asymmetry in the kinds of judgment attributed to conscience (seen earlier): conscience does not judge whether an agent's actions are meritorious.<sup>25</sup>

Although Kant draws a parallel between conscience and a court of law, there is no "statute of limitations" on the judgments issued by conscience, nor is there anything to prevent an agent from feeling the bites of conscience for the same misdeed on more than one occasion. Conscience might judge an agent guilty of having infringed on the moral law for as long as the agent retains a memory of the action. Evidence for this can be found throughout Kant's corpus. For example, in the only mention of conscience in the *Critique of Practical Reason*, Kant suggests that conscientious repentance for a misdeed might

<sup>23</sup> To be fair, there are a number of handwritten notes that suggest otherwise. For example, in a note assigned to the period 1764–1768 Kant defines conscience as "the moral feeling applied to one's own actions" (HN, AA 20: 168.10–11), and in a note assigned to the period 1776–1778 Kant writes that the satisfaction or regret one feels regarding the morality of one's action, if felt "before the action," may be called "feeling," but "after the action [it may be called] conscience" (RefI, AA 19: 266.02).

<sup>24</sup> MS, AA 06: 440.30–32.

<sup>25</sup> Despite Kant's claims in his published works that conscience only punishes or reproaches agents for wrongful action, there are a number of notes in which Kant suggests that conscience might have a more positive role to play. For example, in a note from 1776–8 Kant writes that "it appears that rewards everything good and punishes everything evil, but in different degrees" (RefI, AA 19: 266.20–21). Similarly, in a note from 1775–8 Kant writes that "there are two kinds of happy states of mind: 1. Peace of mind or satisfaction (good conscience); 2. The perennially joyous heart" (RefI, AA 15: 260.22–24). He then goes on to say that the first consists in the consciousness that one has behaved in a morally upright fashion, whereas the second is a gift of nature. This has implications for Kant's account of happiness, a topic I discuss in part 3 of Kahn (2019b).

extend long past its occurrence.<sup>26</sup> In retelling the biblical story of Job in the “Miscarriage,” Kant has Job declare that “his conscience does not scold him at all for the sake of his entire life.”<sup>27</sup> In the *Religion*, Kant repeatedly discusses agents on their deathbeds who engage the faculty of conscience to reflect on misdeeds that might have occurred at any point in their lives.<sup>28</sup> And in the first half of the *Metaphysics of Morals*, Kant asserts that requiring officials to swear an oath at the end of every year (saying that they have fulfilled their official duties faithfully) “would bring conscience more into motion” than requiring them to swear an oath at the beginning of every year (saying that they will do so), reasoning that in the former there will be a cumulative effect associated with failures that will be lacking in the latter.<sup>29</sup>

Kant also thinks that the judgment of conscience might come before an agent’s decision to engage in wrongdoing. In the *Metaphysics of Morals*, Kant calls this a “warning conscience.”<sup>30</sup> This preventive role of conscience is not exactly in keeping with Kant’s court metaphor if one thinks that courts require an actual criminal offense (the *actus reus* in the common law tradition). However, that is not an objection to Kant. It merely indicates one of the limits of the court metaphor. Further evidence for ascribing these ideas about the timing of the judgment of conscience to Kant may be found in the students’ notes from Kant’s lectures, where one finds frequent references to conscience as judging before, during, and after an action, often conjoined with claims about the intensity of this judgment. For example, in one set of notes it is asserted that conscience is most powerful after an action and least powerful during an action.<sup>31</sup>

Thus far I have focused on conscience insofar as it judges whether an agent has infringed on the moral law. However, that is not the only judgment assigned to conscience on Kant’s account. Conscience also judges whether an agent has undertaken to determine what his/her duty is with all due diligence. This comes out most clearly in the *Religion*. In the *Religion*, Kant gives two definitions of

<sup>26</sup> KpV, AA 05: 098.14–32.    <sup>27</sup> MpVT, AA 08: 265.21–22.

<sup>28</sup> RGV, AA 06: 070.08–14 and 078.27–34.

<sup>29</sup> MS, AA 06: 305.08–09. Kant also thinks that if an agent makes a promise at the beginning of the year, s/he is likely to excuse him/herself later on account of unforeseen difficulties, an excuse that will not be available in retrospect. All of this, of course, rests on generalizations from empirical psychology.

<sup>30</sup> MS, AA 06: 440.11.

<sup>31</sup> V-PP/Herder, AA 27: 043.14–16. However, sometimes it is asserted that conscience is least powerful before an action, of middle power during an action, and most powerful after an action (V-Mo/Collins, AA 27: 356.11–18). And occasionally it is asserted that conscience really only judges at one or another of these times (V-MS/Vigil, AA 27: 617.29–31 seems to suggest that conscience does not judge during an action). So the lecture notes should be handled carefully here.



conscience.<sup>32</sup> According to the first definition, “conscience is a consciousness which is for itself a duty.”<sup>33</sup> According to the second definition, conscience is “the moral power of judgment directed toward itself.”<sup>34</sup> To explain his first definition, Kant then argues that (1) “it is a moral principle which requires no proof [that] one should *risk nothing on the danger that it is wrongful*”; thus (2) “of an action that I will to undertake I must not alone judge and opine but rather be certain that it is not wrongful,” something Kant calls a “postulate of conscience.”<sup>35</sup> To explain his second definition, Kant argues that in conscience, “reason is directed toward itself”; in conscience, reason judges whether it “actually has undertaken the former evaluation of actions with all diligence (whether they [the actions] are right or wrong).”<sup>36</sup> The idea is that it is morally obligatory that an agent be certain that his/her action, in any given case, is permissible, and the faculty of conscience judges whether an agent has fulfilled this obligation with due diligence. Kant calls this moral obligation for certainty the duty of conscience. To put this another way, conscience judges not only whether an agent has fulfilled his/her duties (the first judgment discussed earlier) but also whether an agent has been negligent in determining what his/her duties are (the due diligence judgment).

Kant illustrates all of this with an example. He imagines an inquisitor who is faced with the task of passing judgment on a heretic. Kant then poses the following question:

now I ask: whether, if he [the inquisitor] sentences him [the heretic] to death, one could say he [the inquisitor] has judged according to his (although erring) conscience, or whether one could give him conscienceless guilt par excellence, whether he might have erred or have done wrong with consciousness.<sup>37</sup>

The issue, according to Kant, is that it is morally certain that taking a life is wrongful, whereas any exemption to this to allow for heretics to be put to death could be based only on revelation and testimony and, as such, “is never apodictically certain.”<sup>38</sup> Kant concludes that such an inquisitor cannot merely have erred; his action must have been conscienceless, performed with consciousness of its wrongfulness, for “he would venture to do something in the danger that it is in the highest degree wrong.”<sup>39</sup> According to Kant, the

<sup>32</sup> However, it should be noted that Kant does not introduce the first as a definition. Rather, much as in section I of the *Groundwork*, when he introduces a second and third proposition in his derivation of the Formula of Universal Law and neglects to say what the first is, Kant introduces his second definition of conscience one paragraph after the first with the remark that “[o]ne could also define conscience thusly” (RGV, AA 06: 186.09–10; emphasis mine).

<sup>33</sup> RGV, AA 06: 185.18–19.   <sup>34</sup> RGV, AA 06: 186.10–11.   <sup>35</sup> RGV, AA 06: 185.23–186.07.

<sup>36</sup> RGV, AA 06: 186.16–18.   <sup>37</sup> RGV, AA 06: 186.24–28.   <sup>38</sup> RGV, AA 06: 187.04.

<sup>39</sup> RGV, AA 06: 187.09–10.



inquisitor cannot have carried out the duty of conscience (the duty for moral certitude), for had he done so, he would have realized the error of his ways.

The *Religion* is not the only text where Kant ascribes this due diligence function to conscience. There are various handwritten notes in which Kant says that an agent ought not to perform any action if s/he is uncertain about its deontic status. Probably the most explicit is from 1783–4:

The supreme principle of conscience is: that nothing is permitted to be done of which the actor is not entirely certain that it (in general) is permitted to be done. We can undertake nothing in the danger of acting wrongfully.<sup>40</sup>

The due diligence function also comes into play in Kant's discussions of the ethics of belief and freedom of conscience.<sup>41</sup> For example, in a draft of the text Kant published in 1798 as *The Conflict of the Faculties*, he says that "what should be counted as religion must be entirely certain . . . for I must answer for it before my conscience."<sup>42</sup> Similarly, in the "Miscarriage" Kant asserts that one commits a wrongful lie if one is negligent in professing to believe something without ascertaining, with all due diligence, that one does:

Someone thus who says to himself (and, what is the same in matters of religion, before God): he believes, without having cast perhaps even only a glance into himself . . . commits not merely the most incongruous . . . but also the most wrongful lie.<sup>43</sup>

From this it may be seen that, on Kant's account, telling an untruth is not necessary for telling a (wrongful) lie.<sup>44</sup> If one claims to believe something without having engaged in the relevant introspection to determine whether this is so, then one has committed a wrongful lie even if one really does believe it.<sup>45</sup> For example, suppose that someone tracks me down at a conference and asks me, politely if incredulously, whether I really believe the thesis of a paper I published a year before. Pressed for time, I hurriedly but unreflectively reply

<sup>40</sup> Refl, AA 18: 579.08–11.

<sup>41</sup> Indeed, the two definitions from the RGV are taken from a section titled "Of the Guiding Thread of Conscience in Matters of Belief" (RGV, AA 06: 185.13).

<sup>42</sup> VASF, AA 23: 436.29–31. Kant's 1783 letter to Moses Mendelssohn is also instructive (Br, AA 10: 347.21–23). In the passage to which this footnote is appended, Kant assumes that religious truths cannot be known with apodictic certainty. This will become important in subsection 2.4.

<sup>43</sup> MpVT, AA 08: 268.22–269.01.

<sup>44</sup> This is borne out in Kant's (in)famous murderer at the door case. In that case, Kant says that telling the murderer that the victim is not at home is wrongful if one does not hold this to be true even if, as a matter of fact, the victim has slipped out of the house (VRML, AA 08: 427.11–20). A similar case can be found in the *Metaphysics of Morals* (MS, AA 06: 431.27–34).

<sup>45</sup> Conversely, Kant thinks that an agent who says something that is literally false has not told a lie if what s/he says is in accordance with his/her best judgment. For a helpful exposition of Kant on lying, see chapter 14 of Wood (2008).

that I do. Finally, suppose that what I said is true and that, if I had reflected sufficiently, I would have realized as much. This, according to Kant, could be an instance of a wrongful lie on account of my lack of introspection. I am going to revisit these ideas about due diligence, introspection, and certainty in subsections 2.3 and 2.4.

Kant's connection between certainty and conscience might be motivated by the fact that the words in German (*gewiss* for certainty, *Gewissen* for conscience) are related.<sup>46</sup> This could help to explain why Kant makes the due diligence function analytic to the concept of conscience (with his *Religion* definitions). The connection also might be motivated by Kant's opposition to probabilism, a doctrine that says that an action is permitted if there is a preponderance of evidence for its permissibility, an evidentiary bar that, on Kant's account, is too low.<sup>47</sup>

However, modern readers might balk at the claim that there is a duty to be certain that one's actions are permissible. Indeed, some might think that such certainty is more frequently conjoined with the most heinous abuses of morality than the alternative. In light of this, it is perhaps worth pointing out that this aspect of Kant's thought is not obviously central to his ethics. For one thing, due diligence need not result in certainty. For another, there are many alternatives between certainty and probabilism: rejection of the latter does not necessitate adoption of the former. I shall return to this in subsection 2.5.

Kant makes different claims about conscience and duty. In some places, Kant claims that there is a duty of conscience. For example, as seen above, Kant's first *Religion* definition of conscience makes the fact that conscience is a duty analytic to the concept of conscience, and he entitles a section of the *Metaphysics of Morals* devoted to a discussion of conscience, "Of the Duty of the Human Being to Himself as the Inborn Judge over Himself."<sup>48</sup> In other places, Kant claims that a duty of conscience is an absurdity. This can be

<sup>46</sup> Paton (1979, 248).

<sup>47</sup> Kant's own definition of probabilism is different from the one to which this note is appended. Kant defines probabilism as asserting that "the mere opinion that an action could well be right is already sufficient to undertake it" (RGV, AA 06: 186.08–09). A slightly different definition may be found in the editorial notes to the *Religion* in the Cambridge Edition of the Works of Immanuel Kant. The editors quote the Salamancan Dominican Bartolomeo de Medina, quoted also in the editorial apparatus to the Academy Edition of the *Religion* (RGV, AA 06: 506n186), who defines probabilism as saying that "it is legitimate to follow a *probable* opinion even if there is an opposite and more probable one" (emphasis mine). Hegel, who also rejects probabilism, defines it differently again: "an action is permissible, and may be done with an easy conscience, if the agent can come up with *any* good reason for it, be it only the authority of one theologian, and even if other theologians are known by the agent to dissent ever so widely from that authority" (Hegel, 2008, 140).

<sup>48</sup> MS, AA 06: 437.29–30. Confusingly, this chapter and its immediate predecessor are both labeled as chapter 2.