

## *Introduction*

### *Rethinking the Kantian Reflective Ideal*

#### 0.1 The Importance of Reflection

There can be no doubt that Kant thought we should be reflective: we ought to care to make up our own minds about how things are and what is worth doing. The reflective person is not blindly driven on by habitual patterns of thought and desire, by the exigencies of tradition and external authority. She is able to ‘step back’ from all of this and assert herself as the master of her own thought. This is a commonplace Enlightenment ideal: Kant was by no means the first to insist on the importance of thinking for oneself, questioning epistemic authority and standing guard against the insidious power of prejudice.<sup>1</sup>

But in Kant, this ideal takes root in a metaphysics that distinguishes the mechanical operations of nature from whatever can be won in the expression of self-determined human reason. Kant understands the great bulk of prejudices (although not, as we will see, the entirety of them) as a tendency towards cognitive passivity, glossing them as the *‘inclination . . . towards the mechanism of reason rather than towards its spontaneity under laws’* (LJ 9:76; tracking RL-2527 [early 1770s], 16:406; see also LD-W 24:738).<sup>2</sup> To make oneself into a properly self-determined cognitive agent – and ultimately into a properly self-determined human being – is an achievement of some kind. When and how is this won? Kant seems to tell us that we must aim for it on the occasion of every judgment. He repeatedly claims that ‘all

<sup>1</sup> Kant and others debated the question of ‘What is enlightenment?’ in the *Berlinische Monatsschrift* and other venues in the 1780s (see Schmidt 1996 and Ciafardone 1990, 321–75 for texts). The topic of enlightenment also figures widely in Kant’s writings, from ethics to anthropology to logic. One of the main sources of his conception of enlightenment is the discussion of prejudice that figured in eighteenth-century logic texts, including G. F. Meier’s (1752) *Auszug aus der Vernunftlehre*, from which Kant lectured over the course of several decades. For historical discussion of enlightenment and the theory of prejudice in the Enlightenment era, see Schneiders (1983); for a focused account of these issues as they figure in Meier and bear on Meier’s influence on Kant, see Pozzo (2005).

<sup>2</sup> On the methodological issues surrounding working with the record of student notes from Kant’s lectures, and regarding *Logik Jäsche* in particular, see §0.4 of this Introduction.

judgments . . . require a *reflection* – if not before the judgment, then ‘at least following critically after it’ (A260–1/B316–7). He speaks even of some such reflection as a matter of ‘duty’ (A263/B319).<sup>3</sup> With this, we can begin to make out that the Kantian reflective ideal might stand in an uncertain relation to moral requirement and virtue.

Many will be inclined to suppose, in light of remarks such as these, that the Kantian reflective ideal is precious, hyper-deliberate and repugnantly moralistic. Versions of this Kantian caricature abound in exegetical and non-exegetical philosophical work, and likewise across work that is both sanguine and sceptical about Kantianism. An example from sceptical quarters provides an apt illustration: ‘[T]here . . . seems to be something wrong with Kant’s ideal of the rational person. This person is always in control. Reason is always holding onto the reins of the soul, ensuring that mental processes are in accord with rational requirements . . . But there is more in life. Being rationally reflective and being rational are not supremely valuable modes of thought and being, but forms of thought and being among others’ (Zangwill 2012, 357).

It is no exaggeration to claim that Kant accords supreme value to being rationally reflective. The error does not lie there. But just what this means, and what the ideally reflective person looks like by Kantian lights has been poorly understood. My overarching aim in this book is to show why the supreme value that Kant accords to being reflective does not yield the common caricature, and to develop an alternative account of the Kantian reflective ideal.

## 0.2 Modelling a Solution

Why has Kant’s conception of reflection been poorly understood? One problem is the complexity of the textual record on reflection, which I canvas in Chapter 1: there are various notions of reflection invoked in a range of different contexts, of varying degrees of technical specificity. I am going to set those complications entirely to one side for now, to focus just on the idea that reflection is a kind of ‘stepping back’ from the immediacy of judgment and action in order to inquire into, and critically assess, its sources or operative principles. There are certain ways of running with this idea that lead to obvious problems.

<sup>3</sup> In the *Amphiboly*, Kant says that anyone who wants to judge about things a priori is subject to a ‘duty’ of ‘transcendental reflection’; I give an account of transcendental reflection in Merritt (2015).

## Introduction

3

Take Kant's claim that all judgments require reflection. If the relevant notion of reflection is some deliberate consideration of the source of one's taking things to be a certain way – and an assessment of whether that source entitles one to judge accordingly – then the requirement seems overly demanding and out of step with what we generally have in mind when we think of what it is for a cognitive state to be justified. As Andrew Chignell puts it, 'Typically . . . the sort of justification we're interested in is a state rather than an activity. A subject's belief that *p* can *be justified*, even if the subject doesn't *do* anything to determine that it is' (2007, 328). Indeed, it is perhaps owing to its apparent implausibility that Kant's claim that *all judgments require reflection* has scarcely figured in the interpretive literature on reflection; and where it is noted (as in Chignell 2007), there seems to be some readiness to pass it off as a slip of the pen.

But Kant's claim is not one-off. It appears not only as cited in both editions of the *Critique of Pure Reason*, but also throughout the various records of Kant's lectures on logic, in his handwritten Nachlass and in *Logik Jäsche*.<sup>4</sup> It also figures (albeit obliquely) in the *Anthropology*, where Kant claims that 'reflection . . . is required' for any cognition – including sensible experience – because cognitions, one and all, 'rest on judgments' (Anth 7:141). Of course, what Kant might have meant when he claimed that *all judgments require reflection* is a difficult question. The aim of Part I of this book is to address that question in full acknowledgement of the problem just raised. By my lights, the seriousness of the problem comes down to this: if we take the claim that *all judgments require reflection* to lie at the heart of Kant's account of reflection, and if we suppose this reflection to be a deliberately undertaken activity of some kind, then we will be hard pressed to accommodate modes of cognitive activity – modes of knowing – that are perfectly well justified, and quite possibly the expression of a certain cognitive excellence, but that are not deliberate in any direct or interesting way, like sensible experience.<sup>5</sup>

Consider next how a similar set of problems might arise for practical judgment, which in Kant's view is itself a determination of the will, and so properly expresses itself in action. Presumably, most of us act unreflectively much of the time: we just carry on and do what it occurs to us to do. We do

<sup>4</sup> See also LJ 9:76 ('we cannot and may not judge about anything without *reflecting*') and LB 24:161 (reflection is 'necessary for any judgment'). These claims are complemented by an overarching conception of prejudice as 'judgment without reflection' that figures widely in the lectures and handwritten Nachlass: see LB (24:168 (twice), also 165, 167), LPh (24:547), WL (24:863), RL-2519 [1760s] (16:403), RL-2534 and RL-2536 [both c. 1776–78], where prejudice is judgment that 'precedes reflection' (16:408). Further discussion of these claims follows in Chapter 1.

<sup>5</sup> McDowell (1994, 2009) is concerned with a problem along these lines.

not step back from the default views that we have about what to do, to consider in each case what its underlying principle is and whether that principle meets some legitimating standard. But this is what the reflective person of Kantian ethics is imagined as doing. This person is widely supposed to have some particular skill at identifying the ‘subjective principles’ – or *maxims* – on which he proposes to act; and he is supposed to be resolute about submitting those principles to the appropriate test. Kantian maxims are commonly interpreted as subjective principles of action specifying, in the first person, to do action of type A in circumstances of type C for end E.<sup>6</sup> To consider the maxim, the agent not only needs consider what he proposes to himself to do; he must also regard the proposed action as an instance of some action-type, which is linked both to some general description of the circumstances in which actions of that type are warranted or permissible or required, as well as to some general characterisation of the end for which such actions may or ought to be performed. So, our reflective moral agent must recognise himself as being in such circumstances and having adopted such ends as warrant the action in question. Necessary (although not sufficient) warrant for an action lies in its moral permissibility. The special test is supposed to check for precisely that – whether the action, determined as the action that it is in light of its maxim, accords with the requirements of morality. The ideally reflective agent is envisaged as someone who most assiduously tests whether he proposes to act on a maxim whose universal adoption he can coherently will.<sup>7</sup>

There are many problems with this picture of the reflective moral agent. First, it is not clear that one’s maxims can be readily identified, as Kant himself points out on occasion.<sup>8</sup> Second, the proper scope of this reflective activity is unclear. Surely (common sense protests) I can act well – my

<sup>6</sup> Kant calls a maxim a ‘subjective principle of action’ (G 4:421n), but what exactly he means by this and how general such a principle must be in order to count as a maxim has been debated and remains a subject of consternation for Kant’s commentators; among the best recent studies of the difficulties of interpreting Kant on maxims is Kitcher (2003). If one takes it that universalisation tests (the so-called ‘CI-procedure’) form the foundation of moral normativity by Kant’s lights, then one will in turn need to commit to a particular view about what the general form of a maxim is. However, I do not assume this view about the foundation of moral normativity in my arguments about Kant on the importance of being reflective. I take maxims to be general practical commitments about what is a reason for doing what; I take it that, for Kant, these commitments are endorsed when we act (whether we step back and explicitly formulate and assess these commitments or not). The maxims that will particularly concern me in the central arguments of this book are the three maxims of healthy understanding, which Kant claims properly govern cognitive conduct.

<sup>7</sup> Brewer (2000, 2002) queries this picture of the reflective moral agent, and considers whether Kant offers the resources to reject it; however, his conclusions are ambivalent.

<sup>8</sup> O’Neill (1998) emphasises this, citing a memorable remark from *Religion*: ‘we cannot observe maxims, we cannot do so unproblematically even in ourselves’ (6:20). The lesson she (rightly, in my view) aims to draw from this is that the cultivation of virtue doesn’t rest chiefly on introspection and

## Introduction

5

actions can be perfectly well warranted, and quite possibly even morally *good* – without going through all of this. Third, there is generally no time to cogitate in this way – to step back from the immediacy of action to identify one's maxim and perform an assessment of its universalisability. Further, in many situations, surely the *right* action, the morally worthy action, will be one that issues as an immediate response to one's simply seeing one's situation in a certain way. These are familiar objections to what is presumed to be the Kantian reflective ideal in ethics. Critics of Kant who lodge these objections very often embrace some form of virtue ethics, and contemporary Kantians who acknowledge the force of these objections have argued that the resources to address them can be drawn from later developments in Kant's ethics, particularly his account of virtue in the *Metaphysics of Morals*.<sup>9</sup>

The turn to virtue among Kantians is part of a larger philosophical trend. In the past half-century or so, there has been a resurgence of interest in the concept of virtue in ethics and epistemology.<sup>10</sup> In both cases, the movement can be described as a broadening of view from isolated episodes of action or belief to the character of the agent who acts or takes things to be a certain way. This broadened view calls for us to recognise that a comprehensive range of capacities and dispositions – including capacities of attention, perception, feeling and desire – is integral to a general outlook oriented towards the relevant moral or epistemic goods. Now, there are many reasons why philosophers have found inquiry along these lines worth pursuing. But within broadly rationalist quarters – where what makes character good or virtuous is that it is appropriately governed by rational principle – making virtue central conceivably provides a kind of buffer against the caricature, at least in its moral guise.

This is because the entire range of capacities and dispositions proper to virtue will be conceived as shaped – or made what they are – by reason. Reason infuses the whole package, which includes capacities that are passive in their operation, such as perception and feeling. The exercise of such

the self-ascription of maxims. Cf. Grenberg (2005, 49–51, 62–64, 97–103), who argues that O'Neill overstates the opacity point, and suggests that moral reflection centrally involves attentiveness to one's own inner life. While many of Grenberg's criticisms of O'Neill are apt, I argue in this book that reflection – and hence, in turn, specifically *moral* reflection – cannot chiefly be an introspective activity understood along such lines.

<sup>9</sup> This approach to Kantian ethics has gathered considerable steam in recent years; consider e.g. two recent edited collections devoted to the issue (Betzler 2008; Jost and Wuerth 2011).

<sup>10</sup> In ethics, the seminal text is Anscombe (1958). Attention to virtue in epistemology came considerably later – beginning with some of the papers collected in Sosa (1991). Not until Montmarquet (1993) and Zagzebski (1996) was virtue epistemology pursued from cues borrowed from virtue ethics, however.

capacities can be recognised as proper to virtue and, as such, no less an expression of the self-determination proper to a rational being than overt efforts of deliberation and inquiry. The virtuous person will not be pictured as excessively deliberate about meeting moral requirement, because it will be recognised that much of the moral work will already be done simply by seeing one's situation in the right light.<sup>11</sup>

I am sanguine about taking a virtue-focused approach to Kant, as my work in this book will attest. But the approach comes with certain risks – not least the danger of making Kant's critical philosophy, arguably the high-water mark of the Enlightenment ethos, into something that it is not. Much contemporary work on virtue draws on Aristotle, but there is little reason to think that Kant thought especially long or hard about him; in fact, Kant's conception of virtue draws more from the Socratic tradition developed by the Stoics, which has exerted relatively little influence over contemporary discussion of virtue.<sup>12</sup> So, we need to be careful about the philosophical assumptions driving any virtue-focused approach to Kant.

Further, while the recent focus of scholarly attention on Kant's conception of moral virtue might help to dismiss the caricature of the reflective moral agent, it is not clear whether (or how) it can address the problems ensuing from the *general* importance that Kant places on being reflective. Consider again Kant's claim that 'all judgments require reflection'. How

<sup>11</sup> This broadly rationalist tradition of virtue ethics draws typically from Aristotle; an important example is McDowell (1979). Herman (1993, 2007) develops a compelling Kantian account of virtue along these (broadly Aristotelian) lines. It should be noted that the development of broadly rationalist virtue ethics has not been uniformly neo-Aristotelian. Murdoch (1971, 36) argues that 'the exercise of our freedom is a small piecemeal business which goes on all the time' through the exercise of attending properly to persons; she presents her position as Platonic in spirit (and, it should be noted, McDowell 1979 suggests it as an influence). Although Murdoch's essays in that volume attack then-contemporary (i.e., mid-twentieth century) Kantianism, she is consistently careful to distinguish her target from Kant himself; indeed, despite superficial appearances otherwise, her own variety of moral rationalism is not so far from Kant by my lights (see Merritt 2017b). Reading Murdoch has influenced my project here to some extent, although I have not attempted to work with Murdoch's writings directly in what follows, and I won't make an explicit case for the closeness that I find. Grenberg (2013, 292) also notes Murdoch as an influence on her work on Kant's conception of virtue, but takes the *attention* required to live well to be directed at the goings-on of one's own inner life (2013, 24 and 159–86), which I think misses the spirit of Murdoch's distinctive notion of attention – although I cannot argue this point here.

<sup>12</sup> Grenberg (2005, 49–51) notes some of the distorting effects of taking cues from Aristotle when interpreting Kant's conception of virtue. While I agree with her remarks that some of the key differences between Kant and the Stoics on virtue turn on differing views of human nature (see Grenberg, 2005, 20–2), I also think that Kant draws more from Stoic ethics than she realises. See Sherman (1997, 99–120) for the beginnings of an account of the relevance of the Stoics for Kant's conception of virtue – although ultimately, she takes Kant to be more deeply allied with Aristotle. I do not track the influence of the Stoics on Kant in this book, but aim to develop this line of inquiry in future work.

might the common complaint that this requirement is overly demanding play out in the practical case? We can find an example in a recent debate between Martin Sticker (2015) and Ido Geiger (2015). For Sticker, universalisation tests are the foundation of moral normativity, and implicitly the fundamental exercise of moral reflection by Kant's lights. Sticker considers the worry that Kant may have an overly demanding view of moral reflection, which he aims to assuage with the suggestion that we need run the test on a maxim only once – after that, we can simply act on the maxim without again stepping back in this way (2015, 982). Geiger replies that Sticker's proposal effectively waives the requirement to be reflective, at least for the most part; what we should do instead, Geiger suggests, is 'make reflection less demanding' (2015, 993–4).

The spirit of Geiger's rejoinder may simply be to point out that by Kant's lights, a life that is lived well can only be reflective through and through – and not solely when we submit maxims to universalisation tests. That, I would endorse. But I cannot accept the assumption that such a picture of a reflective life should show reflection to be less demanding than the maxim-universalising view. For if we interpret Kant's claim that *all judgments require reflection* as calling for stepping back on the occasion of every judgment, and then baulk at the implausibility of this, we will have already conceded too much to a picture of the Kantian reflective ideal that I aim to reject: we will assume that the requirement is overly demanding, when we should worry that the interpretation of the demand has gone awry.<sup>13</sup> There is important foundational work to be done on what Kant takes *reflection* to be – in general terms – so that we might, down the road, arrive at a more stable and compelling account of its role in moral life. My arguments in this book follow that trajectory.

### 0.3 Précis

In Chapter 1, I begin by drawing a distinction between constitutive and normative requirements to reflect. It is partly constitutive of what it is to possess a rational mind that one has an at least tacit handle on oneself as the source of a point of view on how things are or what is worth doing. We cannot think at all without this; this self-consciousness – or *reflection*, I argue – is a constitutive requirement on thought. Reflection in this sense needs to be distinguished from the consideration of whether one has reason to take it that  $p$  or to  $\phi$ : such questions can be settled in judgment, but there

<sup>13</sup> I thank Bridget Clarke for pressing me to clarify my point along these lines.



is a subjective orientation to such thinking that consists in taking a certain interest in oneself as the one who settles the question. I argue that this is the sense of reflection Kant has in mind when he claims that *all judgments require reflection*. As we learn through close examination of Kant's views on prejudice, his idea is not that it is impossible to make use of one's cognitive capacities at all without this reflection, but only that it is impossible to do so *well*. That is why reflection, in this sense, is a normative requirement on judgment.

However, the account I offer of the normative requirement to reflect in Chapter 1 is only preliminary, as it does not provide ready resources to meet the objection, already raised in this Introduction, that it is overly demanding (or, rather, makes the wrong demands). In Chapter 2, I suggest that Kant offers a more nuanced account of the requirement in question when he formulates three 'maxims' of 'healthy human understanding' in some of his later work. One of my aims is to show that the requirement issued in the claim that *all judgments require reflection* is both normative (we cannot make *good use* of our cognitive capacities without it) and yet need not be conceived as a deliberately undertaken activity of some kind. To that end, I argue that reflection, in this sense, is internal to sound judgment: it is nothing separate from considering the objective cognitive question *in the right spirit*, or *with the right frame of mind*. This is how I argue that the requirement to reflect in this sense lodges at the level of character, rather than piecemeal on the occasion of each and every act of judgment.

In Chapter 3, I take on questions about the relation between the constitutive and normative requirements to reflect, arguing that as soon as the first is met (and thus, there is genuine thought), the latter must be met *to some degree* as well. My aim here is to clarify what is basic to cognitive agency by Kant's lights. I do this by looking into Kant's remarks about perception, attention and experience in the *Anthropology* and in related passages of the *Critique of Pure Reason*. I argue that experience requires attention by Kant's lights; this in turn allows us to understand how the enjoyment of experience is an engagement of cognitive agency, despite its putatively passive character. From this, we can begin to understand how, by Kant's lights, sensible experience is in principle no less the expression of our rational self-determination than overt efforts of deliberation and the like. This concludes Part I, which focuses on the interpretation of Kant on reflection.

The account of the normative requirement to reflect in Part I raises questions about the relation between cognitive and moral character. In Part II, I argue for the thesis that moral virtue is a specification of general cognitive



virtue, and that general cognitive virtue is nothing other than the notion of healthy understanding discussed in Chapter 2: I call this the *specification thesis*.<sup>14</sup> The specification thesis presupposes a certain conception of reason: namely, that reason is at bottom a cognitive capacity, albeit one admitting of distinct theoretical and practical employments. However, some Kantians think that only the theoretical exercise of reason is genuinely cognitive, and assume that when Kant speaks of ‘practical cognition’ – as he often does – the cognition in question does not share anything basic, qua cognition, with theoretical cognition. I disagree: the textual evidence, as I see it, overwhelmingly supports the ascription of the former view to Kant. Since this remains a contested issue among Kantians, and since the specification thesis might seem to some to run afoul of Kant’s remarks about the ‘primacy of practical reason’, Chapter 4 adduces the textual evidence for the conception of reason I attribute to Kant, and explains why my thesis does not get into trouble over the ‘primacy of practical reason’. This sets the stage for Chapter 5, which argues for the specification thesis. There, I argue that healthy understanding is a conception of good cognitive *character*, which I then locate in relation to good moral character through the account of virtue in the *Metaphysics of Morals*. This work underwrites the project that occupies me for the final two chapters, which is to elaborate on the cognitive basis of moral virtue by Kant’s lights.

In Chapter 6, I examine Kant’s qualified endorsement of the idea that moral virtue may be a certain sort of skill (*Metaphysics of Morals* 6:383–4). Exploring the historical context of this remark, and carefully working out its philosophical implications, allows me to begin to make clearer and more determinate sense of the cognitive basis of moral virtue. This is also where my alternative sketch of the Kantian reflective ideal begins to take shape, firmly planted at considerable distance from its widely peddled caricature. Chapter 7 elaborates on the cognitivist implications of the skill model of

<sup>14</sup> ‘Cognitive virtue’ is not a term Kant himself used, and so I should set out with at least a rough and ready account of what I do and do not mean by it. I have chosen not to use the terms most widely in use in contemporary virtue epistemology: ‘intellectual virtue’ and ‘epistemic virtue’. What is meant by these of course varies from theory to theory, but there are two assumptions that may be explicitly or implicitly bound up with their use – or interpretation – in contemporary circles. One is the common assumption that knowing is essentially or exclusively theoretical (i.e. concerns natural or historical facts); the other is the somewhat less common assumption that knowledge is essentially realised only in the explicit grasp of claims and principles. Kant himself assumed neither. As to the first, see Chapter 4. As to the second, we will see that Kant considers at some length modes of knowledge that are possible without explicit grasp of the principles that makes the knowledge in question possible (Chapters 2 and 6). So, I have chosen to speak of ‘cognitive’ virtue to distance myself from either assumption, regardless of the extent to which they may or may not be operative in any given contemporary conception of intellectual or epistemic virtue.

moral virtue, and demonstrates how this model plays out at greater length in the Doctrine of Virtue of the *Metaphysics of Morals*.

#### 0.4 Comments on Methodology

One of my motivations for working on Kant's conception of reflection is to reconstruct Kantian commitments about mental agency. There is a tendency among commentators, when giving an account of the core arguments of Kant's critical philosophy – above all in the *Critique of Pure Reason* – to craft explanations of cognitive activity in terms of what faculty contributes what to the production of knowledge. But it seems to me that we should never lose sight of the fact that it is a *person* who knows, believes, perceives, is inclined to think one thing, judges another. Although there is a place for considering how Kant assigns various cognitive tasks to various cognitive faculties, in my view the core arguments of Kant's critical philosophy should be interpreted in a manner that tethers these arguments to a ground-level view of our cognitive lives, and the nature and scope of the agency that we have in them. In this book, I am mostly interested in the ground-level view; and to maintain some kind of focus on it, it will be necessary to take the results of the core arguments of the critical philosophy more or less for granted. Thus, I will have little or nothing to say about how Kant arrives at the particular set of principles he claims are constitutive of human reason in its theoretical employment, nor about how he stands to claim that the categorical imperative is constitutive of human reason in its practical employment. I am interested, rather, in what follows about the agency of creatures who are so constituted, in some sense, by nature.

Much of my work in this book connects the dots between claims Kant made in various places and in disparate contexts in order to work out his commitments on the topics of interest: reflection and cognitive virtue. This interpretive work takes place where various lines of philosophical inquiry converge – particularly in Kant's ethics, anthropology and logic. While my discussion has roots in Kant's critical-period works, it is mostly in his later works – chiefly the *Metaphysics of Morals* (1797) and *Anthropology from a Pragmatic Point of View* (1798) – where the key ideas that I am concerned with are developed. I have also found Kant's handwritten Nachlass and the records of notes from his lectures on these subjects to be helpful in this reconstructive work. My principles in working with such materials are as follows. First, while there are a few places where I consider at some length this or that remark from the handwritten Nachlass, my intention is that this should only corroborate a picture that rests on an interpretation of the

## Introduction

II

texts Kant wrote and prepared for publication himself. Second, since the handwritten notes – collected in volumes 15 through 19 of the Academy edition as *Reflexionen* on various topics (anthropology, logic, metaphysics and moral philosophy, respectively) – came from Kant's own pen, I tend to accord them a somewhat higher status, as a source for working out Kant's views on a given matter, than the student notes from his lectures on these topics. For we in fact know relatively little about how the lecture notes originated; and in many cases, the notes were likely taken by professional note-takers who may not themselves have had any first-hand understanding of the topics being discussed.<sup>15</sup> This is not to say that the lecture notes cannot inform an interpretation of Kant, only that we should be careful about how we put them to use: they need to fill out and corroborate a picture that is formed by close study of the works that Kant wrote himself, and ideally also saw to publication.

Special concerns hold for the *Jäsche Logic*, which (perhaps owing to its placement in the subset of volumes in the Academy edition devoted to works published in Kant's lifetime) is often treated by commentators as if it were on par with works Kant wrote himself and saw to publication. Towards the end of his life, Kant commissioned Gottlob Benjamin Jäsche to draw up a text of his logic lectures; to this end, he provided Jäsche with his own heavily annotated copy of the logic textbook from which he had lectured over many decades, Georg Friedrich Meier's *Auszug aus der Vernunftlehre*. Kant's notes were written in the margins and between the lines of the text itself, and on interleaving pieces of paper; they are collected as *Reflexionen zur Logik* in volume 16 of the Academy edition. To generate his text, Jäsche can only have interpolated from those notes, and probably also from copies of lecture notes in circulation in Königsberg at the time. There is, further, no evidence that Kant approved the text that Jäsche came up with.<sup>16</sup> So even though Kant commissioned the *Jäsche Logic*, and even though it was published in his lifetime, we have good reason to handle it cautiously. When working with it, I typically begin by checking to see if the remark that I am interested in can be traced directly to Kant's handwritten notes, and cite both in conjunction when such correspondence can be found (noting that the relevant passage in *Jäsche* 'tracks' a given *Reflexion*). Then, at least, I know that the remark is not merely Jäsche's interpolation.

<sup>15</sup> For documentation of some of these issues, see Naragon (2006) and Boswell (1988). For a proposal on how to work with the lecture notes on logic, see Lu-Adler (2015).

<sup>16</sup> As Young (1992, xvi–xviii) and Naragon (2006) both point out. Something similar holds for Friedrich Theodor Rink's compilation of Kant's *Lectures on Pedagogy* (Päd) – although in that case, we know even less, since we don't have the handwritten notes that Kant supplied to Rink.