1 Introduction

The study of belief and its norms of rationality is a central part of contemporary epistemology. But belief is just one of many doxastic states. When it comes to the class of doxastic states, epistemologists commonly distinguish between our *outright* doxastic states and our *degreed* doxastic states. The outright doxastic states include believing that *p*, thinking that *p*, having the opinion that *p*, being sure that *p*, being certain that *p*, and doubting that *p*. Our degreed doxastic states include degrees of confidence, credences, and certain degreed phenomenal states.

However, in addition to the outright doxastic states mentioned above we also have **conviction**, that is, the state of being convinced (*simpliciter*) that something is the case. And in addition to the degreed states mentioned in the previous paragraph we have **degrees of conviction**, that is, being *more or less* convinced that something is the case. The concept of conviction was central to Kant's way of thinking about our doxastic states. However, conviction has not been regarded as a distinctive doxastic mental state in recent philosophy of mind and epistemology. The aim of this Element is to locate and defend the distinctive place of conviction and its degrees among our doxastic states.

When it comes to our doxastic states there are two kinds of questions we can ask. We can ask questions about their nature:

> **Nature Question**. For any agent *S* and doxastic state *D*, what is it for *S* to be in state *D*?

But we can also ask questions about their structure:

> **Structural Question**. For any doxastic states *D1* . . . *Dn*, how are *D1* . . . *Dn* related to each other?

Section 2 begins with a suggestive Kantian answer to the Structural Question. It then provides evidence for a version of a Kantian picture on which we have at least three outright doxastic states, where thinking is the logically weakest state, certainty is the logically strongest state, and conviction stands between them. A version of Foley's (1992) reductive Lockean approach to our outright doxastic states is considered. On this view, we can account for all our outright doxastic states in terms of *confidence thresholds*. This view is rejected, owing to the psychological possibility of having a very high degree of confidence in *p* while failing to believe, or think, or be convinced that *p*.

Section 3 provides an alternative view. It demonstrates the foundations for thinking that conviction comes in degrees and shows how degrees of conviction provide what is needed for a distinctive Kantian Threshold View of our outright

doxastic states. For some readers, the Kantian Threshold View will not appear very different from its Lockean counterpart. This is likely owed to the following presupposition:

> **Conviction–Confidence Identity**. Degrees of conviction just are degrees of confidence.

But this presupposition is plagued with problems. However, to appreciate these problems we first need to answer the Nature Question in regard to degrees of conviction. Section 4 does this, arguing that one's degree of conviction in *p* is, roughly, the strength of one's disposition to rely on *p*.

Section 5 defends the *sui generity* of degrees of conviction. In particular, this section explains how and why degrees of conviction separate from degrees of confidence (credences) and other degreed doxastic states, including felt degrees of confidence, the feeling of conviction, and degrees of revisability. It also provides an ecumenical suggestion about how best to understand talk of 'degrees of belief'.

Section 6 uses facts about masking dispositions to explain how and why we can *simultaneously* believe (/think, /be convinced, /be certain) that *p* while also suspending these very states. This is a significant result as it's usually assumed that suspending an attitude necessarily involves *lacking* that attitude; that is, believing that *p* and suspending belief that *p* are incompatible. This incompatibilist idea is central to many epistemic problems and has been used to motivate dilemmas of rationality. But if belief and the suspension of belief are compatible states, then once-paradoxical cases arguably cease to be paradoxical.

Section 7 turns to historical questions about the extent to which Kant was himself a 'Kantian' in our sense. It turns out that Kant's theory of doxastic states was surprisingly Kantian as he prominently discusses conviction *simpliciter* and occasionally comments on degrees of conviction. Further, there is some evidence that Kant thought about states such as opinion and certainty in terms of degrees of conviction. Lastly, Kant's concept of degrees of conviction is open to (or at least not in tension with) the dispositional analysis of degrees of conviction.

2 Conviction and Its Doxastic Neighbourhood

Belief is the paradigmatic outright doxastic state, and many studies of our outright doxastic states start with belief. This study is different. It begins with a summary of Kant's views about the outright doxastic states, which motivates an exploration of a body of linguistic evidence for the idea that we have at least

three distinct outright doxastic states that involve taking a positive stance towards a proposition: thinking, conviction, and certainty. Reflection on Kant's theory of assent motivates the idea that these are strength-ordered in the following way: thinking is the logically weakest state (entailing none of the others), certainty the logically strongest state (entailing all of the others), and conviction stands between (entailing only itself and thinking). We explain why a Lockean Threshold View cannot account for these structural facts, and in the following section we introduce a Kantian Threshold View.

2.1 A Kantian Approach to the Doxastic Attitudes

This section quickly introduces one interpretation of Kant's understanding of our fundamental doxastic states that stands out to us for the simple reason that it provides the starting point for a promising way of approaching the nature of the outright doxastic states and their relation to degreed doxastic states in philosophy of mind and epistemology.

When it comes to our doxastic states Kant introduces a central and organizing doxastic concept that he calls *Fürwahrhalten*, which literally means 'holding to be true', but is more commonly translated as 'assent'. **Assent**, as a genus, should be thought of as a purely doxastic state and one that sits atop a gothic taxonomy with many species and subspecies that are individuated by various further characteristics (A820–31/B848–59).[1] To begin to get a grip on this taxonomy, consider any arbitrary case in which an agent assents to *p*. Kant suggests that of any case of assent we can ask at least the following questions:

> **Normative Questions**. Is the agent's state of assent justified by the agent's grounds/reasons to any degree? If so, is there a *sufficient* degree of justification for being in that state? Is the justification in question provided by evidential and/or non-evidential grounds/reasons?[2]
>
> **Voluntaristic Questions**. Did the agent come to assent to *p* just by choosing to do so? That is, is the agent's state of assent within their direct voluntary control?[3]

[1] German translations of English sentences are due to the collective effort of the 2023 CONCEPT research group at the University of Cologne, with Chris Benzenberg and Lena Ghareh Baghery shouldering most of the load. The historical scholarship contained in this volume and the translations of Kant are due to Chris Benzenberg. We cite the *Critique of Pure Reason* by the A edition (1781) and B edition (1787). Other references to Kant are to the volume and page of the Academy edition. Reflections are also cited by their R number. For reconstructions of Kant's taxonomy of assent, see Stevenson (2003), Chignell (2007a,2007b), and Pasternack (2014).

[2] Kant's terms for evidential/non-evidential justification are 'objective grounds' and 'subjective/practical grounds'. Whether *S*'s grounds are objective or subjective depends, in part, on whether the grounds for *S*'s assent are universally valid or only valid for *S* (A820–1/B848–9; 24:150).

[3] For Kant, an answer to this question can be implied by an answer to the previous question because a state of assent is within one's direct voluntary control just in case it doesn't rest on sufficient

Doxastic Question. How strong is the agent's state of assent? That is, *how strongly* does the agent take *p* to be true?

Kant derives different species of assent based on how these questions get answered. For example, knowledge (*Wissen*) is a species of assent for Kant that requires assenting to *p* in response to sufficient evidential grounds, where these grounds force one to assent to *p* and thus put one's assent outside of one's direct voluntary control.[4] And knowledge, for Kant, additionally involves the strongest degree of assent. Faith (*Glaube*) is another species of assent for Kant, which requires assenting to *p* on sufficient non-evidential, practical grounds. Unlike knowledge, faith is within one's direct voluntary control. But, like knowledge, faith for Kant requires the strongest degree of assent.[5]

However, while most species of assent respond to all three questions, Kant occasionally also identifies instances of assent that *only* respond to the Doxastic Question. These species of assent involve purely doxastic attitudes that can be characterized by differences in their strength. Put differently, Kant seems to have recognized that agents can assent more or less strongly, and that this feature of a state of assent can be described independently of its normative and voluntaristic features. These doxastic states have been largely ignored in the recent literature on Kant's theory of assent, mainly because they don't feature prominently in the standard taxonomy. Yet they are the ones we'll be focusing on.

Kant identified three kinds of assent in relation to how strong one's state of assent could be. First among these is **certainty (*Gewissheit*)** or more precisely what Kant calls 'subjective certainty' (24:437). Subjective certainty arguably aligns with what we now call 'psychological certainty' and indicates the strongest state of assent.[6] Next is Kant's notion of **conviction (*Überzeugung*)**, specifically his notion of 'subjective conviction' (A824/B852). This consists of

objective grounds (Benzenberg forthcoming). Since the voluntarist profile of assent aligns with its normative profile for Kant, it cannot be used to derive new species of assent that we could not have already derived via the normative profile.

[4] See (A822/B850), (R2507, 16:397–8), (24:158), among many other passages.

[5] See especially (A823–9/B851–7), (5:472), and (R2462, 16:380).

[6] See (R2450, 16:373), (R2459, 16:378), and (R5645, 18:291). This psychological sense of 'subjective certainty' is not to be confused with a normative sense of the same term. Kant sometimes also uses 'subjective certainty' to denote the practical-moral certainty of faith (R6099, 18:452). Moral certainty results from sufficient subjective grounds, which in the case of faith, point to a special kind of subjective-practical *justification* (A823/B851). Note also that for Kant, not all certainty is subjective certainty; Kant's notion of 'logical certainty', for example, roughly corresponds to what many today call 'epistemic certainty' (A822/B850; A829/B857). More on this in Section 7.1.

a strong – Kant uses the term 'firm' (A824/B852) – state of assent. Weaker than both of them is **opinion (*Meinung*)**. Opinion is a bit tricky because Kant also defines it with reference to the Normative Question and its degree of evidential support (A822/B850). But insofar as the strength of opinion is proportional to its evidence,[7] we can bracket this normative dimension and focus entirely on opinion as a weak state of assent, which one might call 'subjective opinion' to keep with Kant's naming scheme. On Kant's account, there is no weaker state of assent than subjective opinion.[8] Since what follows is just about our doxastic states, we will drop the 'subjective' qualifier. Putting a picture to these strength-ordered states of assent, we have Figure 1.

There is much more to say about this interpretation of Kant's theory of assent and its relation to contemporary ways of thinking about our doxastic states. We'll return to this in Section 7. For now these quick remarks on Kant inspire new ways of approaching the structural relations that obtain among our doxastic states. First of all, it is somewhat uncommon for epistemologists to theorize about opinion, but opinions are among our doxastic states. Second, it is especially uncommon for epistemologists to theorize about conviction, that is, the state of being convinced. But it too is among our doxastic states. Third, it is somewhat uncommon to seek to organize our outright doxastic states as standing in something like a genus–species relation (or, at the very least, a generality relation), where the most general outright state entailed by all the other more specific states involves the idea of 'holding a proposition as true'. But in what follows we will argue that this broadly Kantian picture of our outright doxastic states is correct.

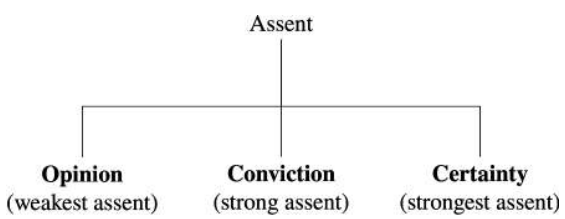

Figure 1 Taxonomy of strength-ordered assent.

[7] See Kant's discussion in the *Logic Blomberg* (24:219). See also Willaschek (2010:189), Pasternack (2011:291,2014:45), and Gava (2019:56–7).

[8] Kant sets the same lower threshold for opinion as he does for assent (24:144; 24:825). Only once, in the third *Critique*, does Kant suggest that there is assent weaker than opinion, namely hypothesis (5:463). But this passage is an outlier that can safely be ignored; after all, Kant typically classifies hypothesis as a species of opinion (24:733). In the same spirit, Chignell notes that Kant classifies 'hunches, working assumptions, and scientific hypotheses' simply as 'weakly-held opinions' (2007b:37). We agree with this reading.

2.2 Natural Language on Thinking, Conviction, and Certainty

Linguistic evidence supports the Kantian idea that our folk theory of mind references at least three distinct outright doxastic states that admit a strength-ordering involving opinion/thinking, conviction, and certainty. This idea runs contrary to an emerging line of thought that natural language (or at least English) makes no reference to any outright doxastic state that stands between thinking and certainty.[9] If correct, there is no outright doxastic state that stands between thinking and certainty in terms of strength.[10] But what has been overlooked is *conviction*. And it is not hard to see that conviction has a distinctive place in our economy of doxastic states.

To begin to see this, take the following expressions, which ascribe outright doxastic states to agents:

(1) *S* thinks that *p*. / It is *S*'s opinion that *p*.
(1′) *S denkt, dass p. / Es ist S' Meinung, dass p.*

(2) *S* is convinced that *p*.
(2′) *S ist (davon) überzeugt, dass p.*

(3) *S* is certain that *p*.
(3′) *S ist (sich) gewiss, dass p.*

Expressions (1)–(3′) are a familiar part of everyday thought and talk. The task to follow is to explore their relational features.

The first thing to highlight is that '*S* thinks that *p*' ('*S denkt, dass p*') and 'It is *S*'s opinion that *p*' ('*Es ist S' Meinung, dass p*') are expressions that seem to refer to the same doxastic state in both English and German. Consider separating them:

(4) ?He thinks that she arrived, but it's not his opinion that she arrived.
(4′) ?*Er denkt, dass sie angekommen ist, aber es ist nicht seine Meinung, dass sie angekommen ist.*

[9] See Hawthorne et al. (2016), Dorst (2019), Rothschild (2020), Holguín (2022), Goodman and Holguín (2023), and Goodman (2023).

[10] In response, some have insightfully argued that we can construct a semi-technical notion of 'strong/outright/full belief' that is stronger than 'believes'/'thinks' in English but also weaker than 'sure'/'certainty' in English. For example, Schulz (2021a) argues that we should theorize with 'outright belief' stipulatively defined as the strongest belief state implied by knowledge. This clever approach has noteworthy limitations. First, there is evidence for weak knowledge; this is implied by Turri's (2015,2016) defence of unreliable knowledge. Second, this stipulative approach leaves open questions about the total class of our outright doxastic states that we seem to refer to with natural language and how they are related to degreed states we seem to refer to. Lastly, there are independent criticisms of semi-technical ways of constructing a concept of 'strong/outright/full belief' in the way that Schulz and others do. See Goodman's (2023) 'The Myth of Full Belief'. See also Heil (2021:ch.4) for a related discussion of Schulz.

(5) ?It's his opinion that she arrived, but he does not think that she arrived.
(5′) ?*Es ist seine Meinung, dass sie angekommen ist, aber er denkt nicht, dass sie angekommen ist.*

These strike us as not only odd-sounding, but also as semantically infelicitous. Should that be so, we will have strong evidence for the following identity:

T=O Necessarily, *S* thinks that *p* iff it is *S*'s opinion that *p*.
Notwendigerweise gilt, S denkt, dass p, genau dann, wenn es S' Meinung ist, dass p.

But not all are happy with T=O. Some informants report that 'It's her opinion that *p*' carries information about the weakness of one's evidential position in regard to *p*. We think this is a pragmatic implicature where the term 'just' is typically heard implicitly and is taken to convey something about the weakness of one's evidential position, as in 'It's *just* her opinion that *p*'. This implicature can also be provoked by adding 'just' to thinking claims, as in 'She *just* thinks that *p*'. Arguably, hearing the silent 'just' is a result of an expectation that speakers conform to the Gricean maxim of quantity: be as informative as one can, and give just as much information as is needed for current conversational aims. For, typically, if one is as informative as is relevant and one is convinced/certain that *p*, one will not indicate only that it's one's opinion that *p*. Further, the sense of evidential weakness associated with opinion comes from the expectation that rational agents proportion their doxastic attitudes to their evidence. Thus, if one just has the opinion (/just thinks) that *p*, this would suggest that one's evidence isn't strong enough for being convinced or certain.[11] We emphasize: nothing turns on T=O in Sections 2, 3, 4, 5, and 6. So we'll sidestep the issue of T=O by privileging 'thinks that' in what follows. We leave it to those who disagree with T=O to say what opinion is and how it differs from thinking-that.

When it comes to discussions of certainty it is standard to distinguish epistemic certainty from psychological certainty, where 'psychological certainty' refers to whatever doxastic state is implicated by the expression '*S* is certain that *p*'. In contrast, 'epistemic certainty' is widely used to refer to whatever condition the expression 'It is certain that *p*' refers to. In what follows we are only concerned with psychological certainty.

It is clearly possible to think that something is true while not being certain that it is. We think a friend will visit today because they said so. But we aren't certain that they will since transportation strikes are not unusual in Cologne. Thus, we have:

[11] In relation to the fact that certainty entails thinking/opinion on our view, one informant noted that they would never say 'It's my opinion that my name is . . . ', but it's something they would say they are certain of. This is also explainable by a Gricean maxim of quantity.

T1 It is possible that *S* thinks that *p*, but *S* is not certain that *p*.
Es ist möglich, dass S denkt, dass p, aber S (sich) nicht gewiss ist, dass p.

Additionally, we may observe that certainty entails thinking. That is:

T2 Necessarily, if *S* is certain that *p*, then *S* thinks that *p*.
Notwendigerweise gilt, wenn S (sich) gewiss ist, dass p, dann denkt S, dass p.

Were T2 false, statements like the following should sound fine:

(6) #I'm certain that he arrived, but I don't think that he arrived.
(6′) *#Ich bin (mir) gewiss, dass er angekommen ist, aber ich denke nicht, dass er angekommen ist.*

(7) #She's certain that he arrived, but she doesn't think that he arrived.
(7′) *#Sie ist (sich) gewiss, dass er angekommen ist, aber sie denkt nicht, dass er angekommen ist.*

But they sound far from fine, and T2 explains their infelicity.

Let's bridge **conviction** and thinking. By 'conviction' we always mean to be referring to *the state of being convinced*. This is not unnatural. For when the noun 'conviction' is used with a sentential complement – 'conviction that' – it tends to be used to refer to the same state that the adjectival expression 'convinced that' does. Some dictionaries inter-define these expressions.[12] It does not matter for the present purposes whether the noun phrase 'conviction that' has other uses in natural language that refer to some other kind of doxastic state. For we will always work with the adjectival expression 'convinced that' and we use 'conviction' as a convenient noun to refer to the state of being convinced.[13]

In regard to conviction and thinking, it is evident that conviction entails thinking:

T3 Necessarily, if *S* is convinced that *p*, then *S* thinks that *p*.
Notwendigerweise gilt, wenn S (davon) überzeugt ist, dass p, dann denkt S, dass p.

[12] For example, the on-line Merriam-Webster dictionary says that conviction is 'the state of being convinced'. www.merriam-webster.com/dictionary/conviction.

[13] We can test the potential equivalence of 'convinced that' and 'conviction that' by considering conjunctions like: 'It is *S*'s conviction that it's wrong to harm innocents, but *S* is not convinced that it's wrong to harm innocents.' This strikes us as odd. Similarly odd is the claim that '*S* is convinced that it's wrong to harm innocents, but it's not *S*'s conviction that it is wrong to harm innocents.' If there is a difference, perhaps it is because there are uses of 'conviction that' which seem to implicate not only being convinced, but also some degree of commitment to remaining so convinced or to somehow promoting that which one is convinced of. But it is not clear to us whether these are semantic implications or pragmatic implicatures.

As evidence for T3 consider claims like:

(8) #I am convinced that he arrived, but I don't think that he arrived.
(8′) *#Ich bin (davon) überzeugt, dass er angekommen ist, aber ich denke nicht, dass er angekommen ist.*

(9) #She is convinced that he arrived, but she doesn't think that he arrived.
(9′) *#Sie ist (davon) überzeugt, dass er angekommen ist, aber sie denkt nicht, dass er angekommen ist.*

These don't make sense. Plausibly, this is because conviction entails thinking.

The last entailment we wish to highlight is that certainty entails conviction:

T4 Necessarily, if *S* is certain that *p*, then *S* is convinced that *p*.
Notwendigerweise gilt, wenn S (sich) gewiss ist, dass p, dann ist S (davon) überzeugt, dass p.

Again, we find evidence for T4 by considering the oddity of instances that run contrary to it:

(10) #I am certain that he arrived, but I am not convinced that he arrived.
(10′) *#Ich bin (mir) gewiss, dass er angekommen ist, aber ich bin nicht (davon) überzeugt, dass er angekommen ist.*

(11) #She is certain that he arrived, but she is not convinced that he arrived.
(11′) *#Sie ist (sich) gewiss, dass er angekommen ist, aber sie ist nicht (davon) überzeugt, dass er angekommen ist.*

The readiest explanation for the contradictory sound of these is T4.

Thinking is not just logically weaker than conviction and certainty (in the sense that it entails neither); it seems normatively weaker (in the sense that it's less evidentially demanding). Let's illustrate this:

Track Race. Take a three-horse race. You know that horse A is more likely to win than horses B and C. You know the probability of A winning is 52%, and of B winning is 24%, and of C winning is 24%. When asked who you think will win, you answer: 'I think horse A will win' – all the while knowing that there is a very good chance that A will not win.[14]

Cases like Track Race are quite ordinary and quite easily constructed:

[14] This is a modification of a case in Hawthorne et al. (2016). I've moved the probability of A winning above 0.5.

> **Diagnosis**. Dr House is treating a patient Sarah in New York with symptoms that are common to five different diseases: A–E. However, House knows that A is not uncommon in New York, but B–E are somewhat uncommon in New York – they are diseases that are usually contracted outside city environments and Sarah rarely travels outside the city. The probabilities of A–E are: Pr(A) = 0.52, Pr(B) = Pr(C) = Pr(D) = Pr(E) = 0.12. When asked to identify Sarah's disease, House answers: 'I think she has disease A.'[15]

Many have argued that one need not speak falsely nor need one manifest any irrationality when claiming that 'I think horse A will win' or 'I think she has disease A' so long as these outcomes are sufficiently more likely than all their competitor outcomes. Indeed, some think that the probability of the accepted outcome can be *less than 0.5*.[16] But here lies a controversy we need not engage. We will limit ourselves to the following lesson:

> **T5** There are at least some cases where it is rational for S to think that p even if S knows that: while p is more likely than $\neg p$, $\neg p$ is almost as likely as p.
> *Es gibt zumindest einige Fälle, in denen es für S rational ist, zu denken, dass p, selbst wenn S weiß, dass p zwar wahrscheinlicher ist als ¬p, aber ¬p fast so wahrscheinlich ist wie p.*

Cases like Track Race and Diagnosis not only provide evidence for the normative weakness of thinking, but also its *logical weakness* relative to conviction. For example, an agent in Track Race might well think that A will win without being convinced that A will win. It would not be incoherent or surprising to hear an agent say in such a case: 'I think that A will win, but I'm not convinced.' This suggests:

> **T6** It is possible that S thinks that p, but S is not convinced that p.
> *Es ist möglich, dass S denkt, dass p, aber S nicht (davon) überzeugt ist, dass p.*

An additional piece of evidence for T6 stems from the phenomenon of *neg-raising* associated with weak mental state terms. When we say 'S doesn't think that p' we tend to provoke the implicature that 'S thinks that not-p'. For example, uses of 'You don't think that she's home' suggest that 'You

[15] Inspiration for this case is from Turri (2015,2016), where he argues for the possibility of unreliable knowledge in cases structurally akin to this. As in Track Race, I have moved the probability of A above 0.5.

[16] See Hawthorne et al. (2016), Dorst (2019), Rothschild (2020), Holguín (2022), and Goodman and Holguín (2023). This idea is noteworthy in the present context since, on Chignell's (2021) reading of Kant, Kant himself seemed to allow for something like 'improbable opinion' (119). See also Chignell (2007a:327,2007b:44) and Section 7.1.