

IS MENTAL PRIVACY DEFENSIBLE?

By

Jaffer Ahmed

Submitted to

Central European University

Department of Philosophy

In partial fulfillment of the requirements for the degree of Master of Arts

Supervisor: Howard Robinson

Vienna, Austria

2021

Abstract

In this thesis I argue that there is a version of the mental privacy hypothesis which is both coherent and plausible despite contemporary objections found in the literature. I argue that the knowledge of mentally private items is direct, non-inferential and a form of acquaintance knowledge, derived from observational and perceptual models of introspection and concerned with the phenomenal character of perceptions and sensations. I show that expressivist accounts and Wittgensteinian attacks against the possibility of knowledge about elements in conscious experience are mistaken and that phenomenal avowals can count as genuine reports, even though phenomenal knowledge is not content-bearing or propositional. I also try to show that knowledge of private mental items involves a substantial but limited form of infallibility and indubitability, and I develop two arguments in favour of first-person authority with respect to knowledge of private items. Finally, I consider whether it is possible to have somebody else's mental experiences and conclude with expressing concerns about a particular argument which is otherwise compelling.

Acknowledgements

I want to express my gratitude to my supervisor Howard Robinson for his kindness, resourcefulness and guidance. I also want to thank the faculty at the Department of Philosophy, particularly Katalin Farkas who played an instrumental role in the genesis of this thesis. Finally, I want to thank my parents Imran Ahmad and Gul Ahmad for their love, constant support and moral encouragement without which the writing of this thesis would not have been possible.

Table of Contents

Abstract.....	1
Introduction.....	4
Chapter 1 – Exposition and Objections.....	9
1. Epistemic and Metaphysical Privacy.....	9
2. Inferential and Non-Inferential Methods of Knowledge	12
3. Wittgensteinian Attacks against the Knowledge Formulation	14
4. Expressivist Account of Avowals.....	19
5. Conclusion.....	21
Chapter 2 - Characterizing Mentally Private Items.....	23
6. Introduction	23
7. Acquaintance knowledge and direct access	24
8. Three Kinds of Epistemic Specialness	26
9. Models of introspection and one argument for first-person authority.....	30
10. Conclusion.....	33
Chapter 3 – Final Considerations	34
11. Can somebody else know/have my thoughts and feelings?	34
12. “Burke’s Assumption”	37
13. Conclusion.....	39
Conclusion.....	40
Bibliography	41

Introduction

Mental privacy is at once the most elusive and the most obvious thesis in modern philosophy of mind. It is elusive because a characterization of it is difficult to formulate, and it is obvious because the idea that each of us knows their minds in a special way which is not available to others has significant intuitive appeal. The issue is also significant because it can result in various interesting philosophical results. For example, if privacy is real, then a kind of skepticism will inevitably arise over whether the mental experiences of different people for which we use the same linguistic expressions can really be known to be the same as well. I can know what I see when I look at things I call green, but I cannot know if you also see the same colour. Similarly, if there is a private realm and knowledge of the contents of this realm is possible, then such a thesis can damage the appeal of behaviorist models of the mind, while strengthening dualistic conceptions. Finally, much of the debate on privacy has been absorbed and subsumed by the knowledge argument controversy, but a defense of mental privacy could potentially show why direct awareness of mental states fails to be captured by physicalist strategies.

The issue has been a part of the philosophical landscape at least since the work of Descartes in the seventeenth century. It was Descartes who first developed the idea of the mind as a private space containing thoughts, feelings and sensations which the person who has the mind cannot doubt (Descartes 1641/1986). Not much later, Locke affirmed that the existence of private space entailed that subjects have an internal sense which they may use to know about the operations of their minds (Locke 1689/1975). Since then, philosophers have traditionally assumed a picture of the mind, one of whose principal features is that it is essentially private, not available for inspection for anyone other than the subject themselves. This conception of the mind has continued to form the background against which contemporary debates in the philosophy of mind have taken place in the last century.

However, modern philosophers have eschewed comprehensive engagement with this fundamental idea in the philosophy of mind, in favour debates which are nevertheless tied to it. The only major philosopher who has written a complete treatment of the subject has been Ayer (Ayer, 1964). His argument has been that the mind is indeed private and that the Cartesian picture of it is mostly correct. Besides Ayer, the present author has not found any treatises or papers which lay out a case for or against privacy by presenting a complete picture in which all aspects of the problem have been discusses. However, there is a vast body of literature about themes which are relevant to mental privacy, although these are rarely discussed with specific reference to privacy. In sections preceding his famous private language argument, Wittgenstein addressed the issue of privacy and concluded that the mind was not private, or at least that our awareness of our minds does not constitute as an instance of knowledge (Wittgenstein, 1953). His arguments centered on the view that knowledge claims can only be legitimately made where the possibility of principled doubt exists, and since one cannot doubt their pain sensations, one cannot know them either. Wittgenstein also developed what has since been called an expressivist account of avowals. According to this view, sentences purportedly reporting mental phenomena are neither truth-apt, nor really statements in the true sense, but only expressions and exclamations. More recently, neo-expressivist accounts have accepted that avowals express self-ascriptions but disagreed over whether they count as instances of knowledge (Finkelstein 2003; Bar-On 2004). The central purpose of these debates has been to assess whether knowledge of mental items is possible at all, and if it is found that it is impossible, the version of mental privacy which is epistemic can be rejected and mental privacy with it. On the other hand, direct perception theorists have suggested that knowledge of mental items is possible through observation of the subject's behaviour but whether this affects mental privacy or not has been debated (McDowell, 1982 and Stout, 2010).

On the other hand, it has also been debated how the items of mental privacy are to be characterized.

Wright has distinguished between phenomenal and attitudinal avowals, and has argued that the latter

are especially insecure (Wright 1998). This has been severely contested as a Wittgensteinian approach that overrides intuitive sense (Mcdowell 1998). David Armstrong has suggested thought experiments to show that knowledge of mental items cannot be epistemically private in principle, since it is at least possible that a machine could be invented in the future that does a brain scan of humans, and reveals the physical characteristics of mental states and their contents (Armstrong 1968). There exists a vast literature on the nature of acquaintance knowledge, which can be used to show that Russell's acquaintance and description dichotomy can prove fruitful in pointing out how knowledge of mentally private items can arise in such a way that includes the doctrine of privileged access (Lewis 1946; Moser 1989; Fales 1996). At the same time, many have criticized the doctrines of self-intimation and transparency which are expected to allow mental privacy living space, presenting various kinds of arguments which appeal to mental obscurity (Shoemaker 1990, Block 1995). Another aspect of the debate that is heavily debated revolves around the supposed infallibility of first-personal mental knowledge, and numerous formulations of this idea have been presented (Shoemaker 1963, Armstrong 1963, Alston 1971, Gallois 1996). Many have argued that subjects are incompetent when it comes to recognizing and understanding their own conscious experience, either due to situational factors or memory issues (Churchland 1988). Similarly, some have construed Freudian ideas as establishing that mental items are not as easily known as imagined, and that subjects are often unaware of aspects of their mental states which are not conscious but still inform and shape the development of conscious experiences. Meanwhile, Chisholm (1981) has defended a limited self-intimation thesis while Jackson (1973) has defended a limited infallibility thesis.

All of these arguments are relevant to mental privacy, because the kind of privacy this thesis focuses on is epistemic, and this maintains that subjects can have knowledge about their mental states but others cannot. But this raises key questions: what kind of knowledge does epistemic privacy require? Is there such a thing as direct access? How does it arise? Does the first person enable to make judgments about

the mind which must always be true? All of these questions and themes can be boiled down to two fundamental questions, which are the research questions this thesis will try to answer. First, what is mental privacy? Second, is mental privacy a defensible idea? Much of this thesis will focus on answering the first question, since the second one cannot be answered without it. My aim in this thesis is to show that there does exist a defensible notion of mental privacy which avoids or successfully counters the many objections found in the literature. It will not be my purpose to develop a full model of mental privacy but only to point in the direction of what a future model might look like.

In the first chapter, I will provide an exposition of mental privacy and show that there are two kinds of privacy involved; mental and metaphysical. Although I will briefly describe their relationship, the rest of the paper will focus on epistemic privacy only. After presenting these formulations, I will argue that privacy requires non-inferential knowledge, and hence privacy may not be defeated simply by appealing to knowledge of other person's mental states via observations of their behaviours. After this, I will turn to Wittgenstein's two arguments against the idea that knowledge of mental items is possible and will argue that both arguments are mistaken. Finally, I will analyze expressivist accounts of avowals, also associated with Wittgenstein and argue that avowals are not mere expressions but also genuine reports. I will conclude that since they can be reported, there is a fact of the matter that can be known, and hence attempts to discard knowledge of private items does not work.

In the second chapter, I will move on to an analysis of characteristics that private items have. I will first try to show that the kind of knowledge epistemic privacy requires is fundamentally phenomenal and a form of acquaintance knowledge which allows for direct access. I will then turn to the issue of infallibility and argue in favour of a limited version of infallibility when applied to judgments about mental events, but a more substantial version of it when applied to their phenomenal character. Finally, I will argue that knowledge of private items depends on an inner sense, and then show how such a model of introspective knowledge does not conflict with limited first-personal authority.

In the third and final chapter, I will briefly consider what prospects an anti-privacy stand can still have. I will show that it is unlikely or impossible for other subjects to have my mental items, and consequently epistemic privacy remains unchallenged. I will argue against both the thesis that it is possible to have other people's mental items and against the view that it is possible to know somebody else's mind. Finally, I will conclude with a short discussion of Burke's Assumption and suggest three concerns that any future development of anti-privacy strategies will have to take account for.

A Note About Structure

Because of a lack of resources directly tackling the issue of privacy, I have had to reconstruct the debate on my own, instead of relying on readings to do that for me. This is reflected in the structure of the thesis, where few authors are quoted across every chapter. My approach has been to identify key themes that have a bearing on the issue, followed by a dig up on readings on those issues in isolation, and then reconnecting them with the privacy debate in ways that are relevant. Consequently, the chapters often contain piecemeal debates and resolutions in the form of short sections into which each chapter is divided. Each of these key themes is a vast subject on its own, so I have opted to include only those debates which serve my purpose. The first two chapters constitute the most serious and weighty part of the discussion. The third and final chapter should be read as a brief consideration of what anti-privacy strategies might still argue for after the main argument of the earlier chapters.

Chapter 1 – Exposition and Objections

1. Epistemic and Metaphysical Privacy

Following Ayer, we may begin by distinguishing between the publicity of matter with the privacy of mind (Ayer, 1964). Objects in the physical world are said to be public insofar as they are as available or detectable to one subject as they are to another. Tables, chairs and persons are all public in this sense. This is not to say, however, that no constraints apply on the detectability of otherwise publicly available objects. We only need to imagine scenarios where someone who is not present in the same room cannot have the same access to its contents as someone who is. Nevertheless, it is easy to see how such items are still similarly available for detection in principle when these constraints are removed. There is nothing in the nature of these objects which makes them more accessible to me and less accessible to you, other than these artificial limitations which can, in any case, be removed if the situation is modified.

On the other hand, the mind seems to work differently, because ostensibly a subject's thoughts, feelings and intentions are not concrete tangible objects in the usual sense laid out along spatial and temporal coordinates. The constraints that apply to these are such that they are irremovable, so that it is logically impossible that anyone other than the subject to know said items. I cannot know whether you are in pain at the present moment, anymore than you can know whether I intend to travel to Berlin the following day. More strictly, to say that a mental object or experience is private is to say that only the subject who experiences it can know it. This is what I shall call epistemic privacy, which is defined as follows;

Epistemic privacy (EP 0): an item in a subject's conscious experience is epistemically private iff 1) the subject knows the contents of his mind and 2) everyone else is necessarily prevented from having similar access to said contents.

This notion of mental privacy entails that subjects are in a unique epistemic position to know what happens in their minds. Each person may know their thoughts, perceptions and sensations but cannot peer into other minds in the same way, and can only speculate or infer what goes on in other people's stream of consciousness. On this account, from the fact that that I cannot know what your sensations are, I can additionally not know whether your conscious experience of pain is the same as mine. I can know that both of us use the word 'green' to describe the colour appearance of the same object, but I cannot know if our visual image is identical. We can follow Raleigh in making this insight the basis of a second formulation of EP as follows;

Epistemic Privacy (EP 1): An element/feature of a subject's conscious experience, *e*, is epistemically private iff it is impossible that another subject can know whether/how *e* is phenomenally similar/dissimilar to elements/features in her own stream of consciousness (Raleigh, 2017).

Both EP0 and EP1 are characterizations of the same concept, the only difference being that EP0 identifies two conditions which a candidate for an epistemically private mental item must fulfill, whereas EP1 expresses how privacy entails the impossibility of conducting a comparative analysis of the mental contents of two different subjects. Consequently, throughout this thesis, I will refer to epistemic privacy simply as EP, although I will have the opportunity to refer to the EP0 as the knowledge formulation in section 1.3.

From EP1, it becomes apparent not only that a mental item is private with regards to the possibility of knowledge, but also ownership or possession. The question of whether two subjects can contrast their

mental contents is intelligible only if it is the case that they both have discreet and separate items in the first place. In other words, the two subjects possess or have their own feelings or sensations which belong only to each of them respectively. This kind of privacy may be called metaphysical, and may be defined as follows,

Metaphysical privacy (MP): an item in a subject's conscious experience is metaphysically private iff it is the subject who has it/experiences it and nobody else.

The formulation of MP above is in need to further clarification, for it can be objected that different people often have the same sensations. Consider the following example. Imagine that there are two subjects who both have pain in their stomachs after having been food poisoned. Wouldn't it be the case that both of them have the same pain? Furthermore, is it also not true that the same may be said of many other kinds of sensations and even perceptual deliverances? In response, it needs to be pointed out that metaphysical privacy does not deny the possibility of qualitatively identical sensations. What it insists on, however, is that in cases such as the one sketched above, even if the subjects are food poisoned in the same way and qualitatively experience the same pain, nevertheless the two pains will be numerically distinct. In other words, there is the pain that the first subject will experience and there is also the pain the second one experiences. Both sensations may be described using the same sensation-word, but their quantity will be two. It is in this way that the two pains will be metaphysically private to their owners.

Although I will be almost exclusively concerned with EP in this paper¹, it is necessary to point out that there is a relationship between EP on the hand and MP on the other. Raleigh has suggested several bizarre scenarios in which EP and MP may come apart. However, putting aside such hypothetical and

¹ Metaphysical Privacy (MP) will be discusses briefly in the third chapter, but is too large a subject to be given full treatment here.

unlikely possibilities, it seems obvious that EP is made possible by MP. A competent, rational and non-Pyrrhonian subject may know the contents of their mind precisely because those contents belong to them. If they were constitutive elements of somebody else's conscious experience, then the subject would not have any epistemic access to them either.

2. Inferential and Non-Inferential Methods of Knowledge

In this section, I will attempt to show why little progress has been made when the notion of epistemic privacy has been defined. Indeed one of the central tasks of this paper is to investigate just what mental privacy is. Although it has become clear that a mental item or object is epistemically private if it is known only to the subject who experiences it, the notion depends on what conception of knowledge it involves and it is to this that I now turn.

Imagine that you are experiencing a certain sensation right now, perhaps a throbbing pain in your hand. If epistemic privacy is to be upheld, then it would seem apparent that only you may know about this sensation. But this is plainly not necessary. It is possible for you to tell another person using ordinary linguistic tools and expressions that you are having this sensation. You may even go so far as describing it and pointing out its location. You may tell them when the pain started and whether it is increasing or decreasing in its intensity. Through all of these ways, the observer to whom you communicate this information will be able to form a more or less reliable understanding of what your sensation is. At the very least, they will be in possession of all the facts pertaining to it and will satisfy the ordinary conditions necessary to make a legitimate knowledge claim. In short, others can know your mental experience because you can transmit all the information that you possess about it, and since testimony is one of the primary sources of knowledge, it seems clear that the observer will know what your sensation is. The problem with this, however, is that it is possible for subjects to lie or make an insincere

report. You may tell them you have a pain when you have none, or you may tell them that the locus of the pain is your hand, when it is in fact your foot. The observer then may be justified in believing what you say, since you are the authority with regards to whatever sensation you have, but their beliefs will be false and therefore not an item of knowledge.

On the other hand, behaviorists and defenders of direct perception models have proposed that it is possible to know the contents of other people's minds by observing their behaviour (McDowell, 1982 and Stout, 2010). This idea has obvious intuitive appeal. After all, in everyday life, we often find ourselves making judgments about the inner states of other people on the basis of the behaviour they exhibit. The constant conjunction of certain mental states with the relevant behavioral expressions ordinarily supplies us with sufficient justification to use the latter to make appropriate ascriptions about one's internal state. When a child cries and groans, we may safely assume that it is in some form of discomfort. When an adult laughs, we can similarly assume that they are amused. Indeed, most agents of reasonable cognitive ability already have a fair understanding of the taxonomy of mental states and their corresponding appropriate expression. However, this kind of behaviorism does not eliminate mental privacy either², because it is not clear why there must always be a one-on-one correspondence between one's sensations and their expression in behavioral terms. It is equally easy to imagine circumstances where a subject may conceal their pain or at least ensure that they do not engage in any appropriate behavioral pattern associated with it. A subject can control their behaviour but it is not immediately obvious if they can do the same with their sensations. When the disproportion between the sensation and its expression is maintained by the deliberate exercise of such control, the epistemic privacy of the sensation will patently fail to be violated.

² It is worth pointing out that although some direct perception theorists think that their work is incompatible with privacy, not everyone does,

Finally, both of the cases discussed above fail to undermine epistemic privacy for an additional reason, namely the fact that both point to ways of acquiring knowledge about the mental contents of other minds through indirect means (testimony in one case, and inference in the other). The kind of knowledge that epistemic privacy is intended to safeguard, however, is not indirect in this way. The key point is that our sensations are supposed to be private because we have direct access to them, which does not require any mediation or inferences. Consequently, whether mental privacy is defensible or not must be answered in terms of direct and unmediated access to the inner states of a subject. This is partly why commentators who are sympathetic to privacy often refer to the contents of such experience as logically private objects, or logically private items. I will return to the question of what directness and logical necessity could mean when applied to non-propositional phenomena such as items in one's stream of consciousness in the second and third chapters.

3. Wittgensteinian Attacks against the Knowledge Formulation

From the previous two sections, we can draw the conclusion that there is some non-inferential knowledge about a subject's mental events which "directly accrues to him through the fact that these thoughts and feelings are his own" (Ayer, 1964). In this section, I will focus on the first of two challenges to this notion, proposed by Wittgenstein and his supporters which threaten to undermine EP by casting doubt on the propriety of applying knowledge terms to the awareness of conscious experience and on the reportorial status of propositions expressing first-personal, subjective experiences. Both strategies are negative, insofar as they involve a rejection of key elements of EP as it is stated in 1.1. However, I will only consider the arguments against the knowledge formulation here and the second set of arguments will be the subject of 1.4.

Wittgensteinian arguments against the knowledge formulation (i.e., EP0) take two broad forms. Both strategies simply deny that the realm of thoughts, feelings and sensations which are a part of conscious

experience can meaningfully be subjected to the application of knowledge terms. This would imply that there is something fundamentally wrong with uttering sentences like “I know that I am in pain”. Taking the first strategy first, it is argued that it is only ever appropriate to say of an agent that they know something if it is at least possible for them to doubt it. Where doubt is not possible, knowledge is not either (Wittgenstein, 1953. Wright 1989). Alternatively, it may be said that it is nonsensical to use the word “know” in relation to propositions which cannot, in principle, be doubted, and that perhaps some other non-cognitivist expression is better suited to be applied to such cases³. From this premise, the argument moves onto the next step where it highlights the fact that sensations such as pain are such that they are impossible to doubt. A subject that experiences intense pain after stubbing their toe cannot sincerely negate or raise questions about whether the pain they are in is real or imagined. Indeed, raising such doubts would be impossible. The impossibility to raise doubts then translates to various other impossibilities, such as the impossibility of giving evidence, investigating, verifying and being corrected. From these premises, the conclusion is drawn that it is false to say that the subject knows what the constitutive components of their conscious experience are.

There are two important things to note about the argument above, which will need to be addressed if privacy is to be maintained. First, it brings to light the indubitability typically associated with private items such as pain sensations. Second, it relies on a principle which puts limits on the range of cases where a legitimate knowledge claim may be made. I will tackle the nature and limits of the supposed indubitability of mental items in the second chapter for an extensive treatment. For now, let's turn to the contention that one cannot know something if the possibility of doubt is excluded. It is better to conceive this principle as a proviso set up by Wittgenstein, rather than a discovered empirical fact about

Snowden (2011) has argued that Wittgenstein’s objection can go away if the words “I know” are replaced by “I realize that”. However, such alternatives also have a cognitivist element, so it seems that Wittgenstein’s point will reapply³

the world. And it is not clear why such a proviso should be accepted. One can ask a defender of the principle to prove its soundness, but Wittgenstein does not offer any evidence for this purpose. More significantly, what does it mean to say in a particular case that doubt is impossible? It cannot mean that doubts are absent because the propositions under consideration are necessary truths, since the laws of logic are necessarily true and yet are genuine examples of knowledge. What Wittgenstein must have in mind are cases where something is not logically necessary but is still beyond doubt. Take the following example. That there is a table in front of me is not logically necessary. There is no logical contradiction in negating view that the table is present, and I can easily conceive of a scenario in which the thing before me was not a table but some other everyday object. At the same time, it is possible for me to doubt the presence of the table, but such doubt would nevertheless not be quite reasonable. Its unreasonableness would be determined by the fact that I can see the table, and visual perception is a strong kind of evidence in its favour. Indeed, it is sometimes the only kind of evidence required. Yet, it is at least conceivable, even if highly unlikely, that I am under a grave misapprehension, or that my senses are deceiving me in some way and leading me to believe that the table is present when it in fact isn't. It seems clear from this example that there is some difference between something being doubtful in principle and something being reasonably doubtful, and that the two can have different outcomes.

What about my pains? Here it seems that it is both unreasonable and impossible to doubt when I am in a painful state that I am suffering in the way that pains affect me. In this way, if I feel pain, then I cannot doubt that I am in pain (the same point can be extended to other kinds of sensations). But from this it follows that I cannot doubt any of my internal states so long as I can have them and feel them. I can doubt whether the visual image I have in my head when I see the table corresponds to reality, but I cannot doubt that I have the visual image when I do. Therefore, if Wittgenstein is right, then I cannot know whether I have a visual image at all.

The real question to ask is this; why is the presence of the possibility of doubt a necessary condition for knowledge? Using a justified true belief account, it is obvious that I can have a true belief about having a visual image and be justified in holding such beliefs. But Wittgenstein's point is precisely that, although the belief might be true, it will lack justification. His reasons for holding on to such a position seem to be that justification requires an independent standard, because there needs to be an objective criterion of correctness without which there will be no way of distinguishing between correct and incorrect beliefs. But feeling pain *is* the objective standard of being correct that one is in pain. It would be absurd to insist that one can feel pain and yet be incorrect about having pain, without maintaining an additional absurd idea, namely that a subject can feel sensations he does not have. If I feel pain, then I must be correct that I have pain and if that is the case, then I can have a true belief with justification. Additionally, another motivation to demand an independent and external standard might be that we must be able to convince each other of our true beliefs by referring to those external facts. However, the very idea of mental privacy is that no such independent standards exist and that we cannot convince each other of them or compare our mental items. In effect, Wittgenstein's proviso is constructed in such a way that it automatically precludes mental phenomena, and then faults mental items for not being able to meet the criterion it has established. The proviso is derived from non-mental states of affairs, where independent and external standards of correctness are entirely applicable, but to then apply the proviso to mental phenomena is to treat mental and non-mental objects in the same way. It will hardly be necessary to show they are not the same kinds of things at all. Non-mental objects are concrete, tangible and material whereas sensations such as pain may or may not have physical sources in the brain, but are nevertheless subjective and intangible items. Based on these considerations, it is safe to reject Wittgenstein's proviso for the purposes of thoughts, feelings and intentions, and to secure the application of cognitivist terms for private items of consciousness.

But there is a second method of undermining the knowledge formulation which does not depend on the proviso but is closely connected to it. This therapeutic reading does not place any necessary conditions on knowledge, but identifies and locates the awkwardness of saying “I know that I am in pain” as part of a grammatical error in the Wittgensteinian sense. The argument is that we can say “I am in pain”, which as is commonly done in ordinary life, and explain pains as something we just have, rather than things that we bear an epistemic relationship to⁴. On this account, pains are not things to be known but things to be had, and we can simply substitute avowals such as “I am in pain”, in the place of other sentences such as “I know that I am in pain”, without any change in either the meaning or truth value of the two statements. However, this is not quite right. That we rarely prefix the expression “I know” to sentences about our mental states is explained better by its unnecessary, since the expression is already assumed to be true. I don’t say that I know my pains, but only because I take it for granted that my knowledge of my mental state is already accepted by my interlocutors, rather than out of a concern for the expression’s use being incorrect. I don’t have to add the prefix, since doing so would be superfluous. Ayer’s proposal that we could theoretically make a list of all items of knowledge that we possess and that such a list would certainly include statements such as “I know that I am in pain”, seems undoubtedly true. If someone were to ask me what all the things I know are, I would not only repeat all the facts in my memory but also that I was in pleasure yesterday and that I am in pain today. Furthermore, such statements also have a truth value. We can imagine a case in which a subject using the expression conveys information which he knows to be false, i.e., a subject can lie. But if the information he conveys is false, then there is something about the case which is true. Hence, the fact that expressions do have a truth value, and that there is a fact of the matter about cases which the

⁴ Although I have opted not present it in the main text, but some of Wittgenstein’s remarks in Blue Book seem to rely use linguistic convention as evidence on its own that avowals with cognitivist prefixes are inappropriate, but it is more charitable to interpret his argument as not appealing to linguistic convention so much as explaining it.

sentence describes is sufficient to demonstrate the legitimacy of expressions that take the form of “I know”, followed by a report of one’s mental state.

4. Expressivist Account of Avowals

In the previous section, we saw how one strategy to undermine, and consequently reject, mental privacy of the epistemic sort centers around denying one of both clauses of knowledge formulation. More specifically, it aims to show that supposedly private items of conscious experience cannot legitimately qualify as instances of knowledge. One of the reasons why some philosophers have objected to the use of expressions such as “I know” when applied to avowals is a particular kind of skepticism about the legitimacy of avowals as genuine reports. In this section, I will undertake an analysis of this line of thought and conclude that avowals can and do obtain a reportorial status.

It is uncontroversially true that expressions such as “I know” or “I believe” may only be meaningfully prefixed before statements which are capable of being either true or false. Exceptions do apply to cases of acquaintance or skill-based knowledge but neither concerns us at present. If there is to be knowledge of mentally private items, which there must be to save the knowledge formulation contained in EP 0, then it is necessary that avowals must express a statement with a content and be examples of genuine reports with a set of appropriate truth conditions. However, the expressivist account offers arguments for an alternative treatment of avowals, where self-ascriptions are understood not as content bearing and knowable propositions, but expressions of the very beliefs that are self-ascribed.

This interpretation of avowals has been dubbed ‘simple expressivism’, and it represents a fundamentally deflationary character of self-ascriptions (Bar-On, 2004). It explains the asymmetries associated with self-ascriptions in the following way. An agent uttering the statement “I believe that p” will only be correct if he does in fact have the belief that p. However, the expressivist account maintains that in making such an utterance, the subject only expresses p, and does not express any beliefs about it

despite the linguistic mislead. The key point is that avowals such as “I am in pain” do not report anything about my mental state, but rather simply express them. On this view, what appear to be reports of mental phenomena are nothing more than a sophisticated expression of it, in the same way as “ouch!” is for pains or a sigh might be for boredom. Just as it would be absurd to claim that a sigh is true or false, similarly it would be equally odd to make such claims for knowing an avowal. “I am in pain” is consequently an expression of pain (just like non-verbal expressions such as a grimace) and not a belief or knowledge claim that the pain is occurrent in the present tense.

There are at least three reasons why I think this account of avowals as mere expressions is gravely mistaken. However, before offering my objections, it must be noted that the argument is not entirely without plausibility. It does appear to be the case that there is no significant difference between saying that one is in pain on the one hand, and that one believes one is hurting on the other. But the first problem with simple expressivist accounts is that the move from apparent report to expression can be inverted to a move from expression to report, based on the exact same line of reasoning. Whatever reasons there are to suppose that avowals are mere expressions are also equally strong reasons to say that expressions are statements. There is nothing in the argument that disallows the construal of exclamations such as “ouch!” as a statement expressing not just a mental state but also a statement that the subject is in pain. This may be illustrated by the example of a person who utters “ouch”, not as an involuntary and spontaneous wince, but as a deliberate attempt to draw attention to and convey information about their being in pain to others. On the expressivist view, such reporting of mental states should not be possible, since there is nothing to an avowal other than its expression. Secondly, the expressivist thesis cannot account for certain kinds of cases. Imagine a student tells you that they are stressed about what they should do to secure an A grade in their upcoming exam. Clearly it would be mistaken to maintain that the student has solely conveyed their anxious feelings to you. They have not done anything like a mere exclamation such as “ouch!” but have also told you what their thoughts are.

Finally, the fact that a subject can lie about what their feelings are at a given moment is evidence that they can also know them. The only which they could not know is if there was no possibility of deception, and since this always exists in ordinary cases, it is safe to conclude that the subject can know their avowals in the relevant sense, and this knowledge can be easily given a verbal and linguistic expression in the form of a statement. This is not to deny however, that avowals perform an expressive function. The criticisms I have offered do not require that avowals must *only report*, but the expressivist thesis does require that they must *only express*. Since this is not true, I conclude that avowals are truth-apt and do have a reportorial status with respect to self-ascribed mental states, in addition to any other functions they may perform.

5. Conclusion

In this chapter, I have presented formulations of two kinds of mental privacy, namely epistemic and metaphysical. I have also tried to illuminate some ways in which the two are related. I have taken both of these formulations as starting points, and not as final definitions. Indeed, it is my purpose in the later chapters to clarify elements of these formulations which are vague or ambiguous at present. After doing this, I have tried to show that the kind of knowledge required by epistemic privacy cannot be inferential and that versions of direct perception theories consequently do not touch the issue. In the third and final section, I have analyzed two kinds of arguments proposed by Wittgenstein and his defenders aimed at undermining the knowledge formulation; 1. That the use of knowledge terms to refer to our awareness of mentally private items is inappropriate and 2. That subjects can have sensations but not know them, since sentences expressing the two are equivalent in meaning. Both of these objections have been found lacking. Finally, I have considered the expressivist account of avowals, and argued that avowals are not only expressions of self-ascribed mental states but also reports with truth conditions

and consequently genuine examples of knowledge. Now that the knowledge formulation has been defended, I will attempt a characterization of mental privacy in the next chapter, and identify its properties and attributes.

Chapter 2 - Characterizing Mentally Private Items

6. Introduction

In the previous chapter, I tried to show that knowledge of mentally private items is possible and that EPO survives the onslaught of Wittgensteinian and Expressivist attacks. It remains to be seen, however, what species of knowledge epistemic privacy involves and how such knowledge arises. In order to answer these questions, it is necessary to attempt a characterization of mentally private items by clarifying both what the properties frequently associated with them mean and the limits within which they must operate. Accordingly, this chapter will contain such an analysis, and then use the findings to present a model of the knowledge of private items in the next chapter.

As a preliminary to the main discussion, let us begin by following Wright's distinction between two kinds of avowals: phenomenal and attitudinal (Wright 1998). Phenomenal avowals are reports and expressions of first personal feelings, including sensations and emotions, whereas attitudinal avowals have some propositional content or intentional direction, and include both thoughts, expectations and intentions. Wright identifies three key features of phenomenal avowals. First, that they are *groundless*, such that it is inappropriate to demand evidence to justify them. Second, they are *authoritative* in that the subject's utterance of them is itself the criterion of their correctness. Third, they are *transparent*, insofar as it would be absurd to say of them that the subject whose stream of conscious experience they appear in can deny knowing them (for example, "I am in pain but don't know it"). On the other hand, attitudinal avowals can be discovered through self-interpretation which in turn lacks all three characteristics. It is reasonable to demand justification of them, depleting their groundlessness. They do not possess authority, because they can be mistaken despite the subject's own views. And finally, they are not transparent, since it is possible that the subject is ignorant of their own intentional psychology.

The contrast between phenomenal and attitudinal avowals has been presented here because it shows that a) not all mental items are the same and because b) it introduces properties associated with mental phenomena. Having introduced these terms, I will now begin the main discussion.

7. Acquaintance knowledge and direct access

In this section, I want to take a closer look at what acquaintance knowledge is, and how it can account for a direct awareness of mentally private items.

David Armstrong presents the case of the neuroscientist who possesses all the physical facts that there are to know about the brain processes that give rise to mental states such as the state of being in pain. He asks us to imagine a point in time in the future when brain science is in such an advanced stage of development that it is not only the case that such facts are fully known but also that machines capable of unveiling such processes have been invented (Armstrong, 1968). Wouldn't it be true that the neuroscientist can scan a subject's brain and thereby know that they are in pain? If we grant the premises, then we must accept that the neuroscientist will indeed know whether the subject is in pain or not as long as the mind-brain identity theory is accepted. However, it is not my purpose here to argue for or against either, especially since there is already a large body of literature devoted to the subject. As an aside, it is not clear how such machines would work. Nevertheless, such machines will fail to recreate the phenomenal character of the pain, which will remain concealed to the neuroscientist, and this is the key point. The neuroscientist may know that the subject is in a painful state, thanks to the advancements of brain science and the development of the relevant technology, but this will not in and of itself be sufficient for him to know what that pain feels like to the subject who has it. At most, he will be able to use some metrics or standards of measurement to describe the physical facts pertaining to the subject's mental state, but we have already seen in the first chapter that the kind of knowledge

involved in epistemic privacy is direct and unmediated, which the machine will still not be able to reproduce. The neuroscientist will be aware of the mental state of the subject, without having the kind of direct introspective awareness of the qualia associated with it. These examples also draw out the insight that the private items of experience may be candidates of two different kinds of knowledge: knowledge about their occurrence in the present tense, or knowledge about their nature or phenomenal character. A defender of mental privacy can claim that epistemic privacy means either that the subject alone knows *that* they are currently having a certain mental experience or that only the subject knows how being in that state at the present time *feels like*.

To say that we are acquainted with our mental items phenomenally is to say that there is a relation of direct awareness in which our mental items are given or presented to us (Lewis 1946; Moser 1989; Fales 1996), which is what Armstrong's neuroscientist lacks. Such direct awareness has two key features; first, it is non-intentional and second, that one can only be directly aware of phenomenal items if those items are in fact present. Mentally private items fulfill both of these criteria. To say that direct awareness is non-intentional is to say that the awareness of mental objects does not involve the making of a judgement. Intuitively there is sufficient appeal in the notion that in being aware of our pains, we encounter the pains first, and any judgments about them are secondary in nature. The failure to distinguish awareness from judgments will lead to a high (rather than moderate) degree of fallibility which I will argue against in the next section. As for the second point, it is obvious that there indeed are such things as mental items such as pain and visual perceptions. This much has never been disputed. However, whereas judgments can typically be expressed in language, the objects of direct experience cannot. We often use metaphors to segment degrees and kinds of pain. A pain might be throbbing, numb, sharp or otherwise but such descriptions are approximations at best, and none can successfully describe the nature of pain currently experienced by a subject. This inability to fully translate phenomenal objects into linguistic expressions thus makes mental items private in a new sense. A

subject can be directly aware of their mental state but since no amount of linguistic contrivance can communicate their phenomenal qualities, other subjects may either only know that the subject is having a pain sensation if the subject makes a report of it (alternatively, Armstrong's neuroscientist may perform his brain scan) or they can have a remote and speculative idea of what the pain might feel like based on the vividness of the metaphors used to make the description.

The acquaintance theory leads to two implications about the privileged access I have to my mind, neither of which can be accepted without qualification. If I am directly aware of my mental states, then this means my mind is self-intimating, and everything that happens in it is completely transparent and open for my inspection (Shoemaker 1990). However, it is possible for a subject to sometimes be cut off from their mind and its contents, or at least for their awareness to be weak and hazy. This can happen when subjects are tired, or inebriated. Subjects also sometimes "zone out" when performing repetitive tasks, and don't fully return to an attentive state of mind until later. Armstrong gives the example of the truck driver who returns to an attentive state after driving for long periods of time, only to realize that he does not remember what he saw or felt during most of the journey (Armstrong 1981). From this, I conclude that our direct awareness of our minds is a capacity, which it is possible in some scenarios to not exercise. Although the subject may continue to have phenomenal experiences, they will not have full and unconditional awareness unless they are attentive and focused on their phenomenal mental objects.

8. Three Kinds of Epistemic Specialness

There are three kinds of epistemic specialness which the knowledge of our private mental contents can have, although their corresponding terms have frequently been used interchangeably in the literature. The first claim is that a subject's knowledge of constitutive components of his conscious experience are

infallible in the sense that the subject's introspective reports about them must necessarily be true. The second claim is that such knowledge is *incorrigible* if other persons cannot have any justification that would induce the subject to revise or alter his beliefs. The third and final claim is that of *indubitability*, which would make it impossible for the subject to doubt or reject his knowledge about his mental state and contents (Shoemaker 1963, Armstrong 1963, Alston 1971, Gallois 1996).

If there is such a thing as mental privacy, then knowledge of our mental contents will be necessarily incorrigible. Others cannot change our understanding of our inner experience when they do not have access to it. We can therefore leave incorrigibility out of the discussion and focus on infallibility and indubitability alone. Quite besides its philosophical soundness, it is intuitively compelling that a subject cannot doubt that they are in pain when they are in pain. One needs to be only reminded of the last time they hurt themselves to grasp that pain is a discomfort intense enough that it overwhelms the subject when it occurs and becomes impossible for the subject to ignore. However, Churchland has pointed out that our introspective judgments frequently go wrong because of memory effects, expectations and presentations (Churchland 1988). I may think I am in pain when I'm being comforted by an ice cube, if I was earlier subjected to torture with iron rods repeatedly. I may not reacquire my ability to distinguish between different kinds of pain and pleasure sensations if I have spent many years deprived of both senses. Similar problems arise when the claim is extended to other kinds of mental contents such as perceptions. Our visual experiences are frequently confused and imprecise. Given these considerations, it would be unreasonable to insist that our knowledge of our inner experiences is always accurate and incapable of error. However, it would still be true that we usually know our mental states well enough, or at any rate, that we know them better than other people do, because we have a direct access to them which others lack. Circumstances which bear on the situation can, and do, intervene on our ability to know what we feel but this is far removed from the contention that we are always or even mostly wrong about it. The same reasons which can warp our senses are also reasons for

the distortion of our knowledge of our inner states. On the contrary, it would be absurd to insist that a subject must know his mental state, *no matter what*. The possibility of error does not undermine privacy, since in order to sustain the knowledge claims we make about our mental states, it is only required that our reports are true, not that they are infallible. Consequently, a defender of mental privacy need not commit to such an extreme position which underplays environmental and contextual factors, and denies that they can ever intervene on the mental.

Consider a person who is currently under a grave misapprehension or hallucination that gives rise to visual images which they perceive, that do not correspond to what is actually before them in space. Suppose also that this person already knows from their past experiences or psychiatric reports that they have a history of hallucinating because of certain imbalances in the brain. Given this, we may ask the following question; in what sense is this subject's visual experience indubitable? It cannot be indubitable in the sense that the efficacy of their mental representational ability to produce visual sense data that reliably conforms to what is actually before them, is beyond doubt. After all, their perceptual deliverance is false and they already know that their illness predisposes them to have visual experiences that are imagined and not real. Still, there is one sense in which their sensory perception may be said to be indubitable, namely that they have such and such visual sense data before them, whether or not it is the way things around them really are. Similarly, although it is possible for the subject who is in pain to doubt whether or not their bodily system is in the right state necessary to produce the pain sensation (such as whether their pain receptors are appropriately functional), they still cannot doubt the feeling of pain which they currently experience in ordinary circumstances. In short, to say that a certain mental experience is indubitable is only to say that there is no doubt that such and such is what it seems or feels like, which can be pointed at using a demonstrative. As such, it is not a thesis about the reliability of the mechanisms that are involved in the production of those experiences. This does not preclude the

possibility however, that we often make false judgements about our inner experiences, but it does rule out the impossibility of being aware that it currently feels *like this* (McGrew 1995& 1999).

I see a friend in the other end of the room, only to find out that I actually saw someone else. Such errors are so commonplace that perhaps it would be wrong to say that even just my phenomenal experience is beyond doubt. But what is erroneous here is not the visual image in the mind but the judgement that I form on the basis of the image. I saw the room and the people in it, and unreflectively arrived at the judgement (however quickly) that the other person was my friend, but this judgement is logically secondary to the image in my mind, since I would not be able to see anyone *as anything* if I did not first have an image. That such judgments are in fact made, however unreflectively, is the source of sufficient confusion which obfuscates the cleavage that separates raw phenomenal viewing and judgement-laden viewing. It is an observed phenomenon that subjects not only see objects in physical space, but that they seem them as being a certain way and the two so often coincide temporally that their logical order of priority is ignored. If I am right however, the division does exist, is purely logical and circumscribes the extent of phenomenal knowledge only to raw feels, and not to Wittgensteinian aspect-seeing.

In considering the different varieties of mental processes (such as perceiving and feeling), we are forced to confront a further problem. In cases such as perception, the subject has a mental image which is ostensibly connected to the external world. When we look at a chair or a table, the mental image we end up having is directed outwards. But sensations don't seem to work in the same way. A pain might be induced by my interaction with physical objects, but both the subject and object of the pain is located inwards, due to which the distinction between the occurrence and feeling of the sensation evaporates. I can see things that are not real, but I cannot feel pains which aren't there. But this need not imply that a minimum standard of privacy cannot be maintained. We can still say that the phenomenal quality of seeing a certain image and feeling a certain pain is indubitable, and then claim that the realness of the

pain is a fact additional to the subjective feeling, that nevertheless coincides with it in toto. So even if the barometer of the existence of pain is that one can feel it, what remains indubitable in both cases is the phenomenal character of conscious experience. Consequently, I cannot doubt that I see p and I also cannot doubt that I feel q, even if I can doubt whether what I see is real and not doubt that I do have a pain, and even if I can make an error in judgment. What I will still possess is acquaintance knowledge with the phenomenal character of mentally private items, which others cannot have.

To summarize, knowledge of our conscious experience is not infallible in the sense that we can never make mistakes about it, given that our internal capacity to understand our inner mental states can be skewed due to the environment and context in which the review or self-scanning is conducted. But we can still have a moderate infallibility and indubitability with respect to the raw phenomenal nature of our mental items, so long as it is non-intentional and logically prior to any associative judgments.

9. Models of introspection and one argument for first-person authority

In the second section, I had argued that we have acquaintance knowledge of our private mental items, since our mental items are non-intentional and are a form of direct awareness sans any mediation. This devolves into an observational model of introspection, according to which our inner sense is analogous to perceptual ability. Just as the latter allow us to see things in the outside world, we have an inner sense which allows us to look inwards, and discover our mental contents and states. Before proceeding to possible criticisms of this view, I want to briefly consider the way in which such direct knowledge arises. It is not the case that we are always aware of our own mental states, because as has been pointed out earlier, awareness sometimes requires a sharp focus which a given person must set on his own mind. On the other hand, it would be misleading to say that one is never aware of one's mental states, since some of them such as pain sensations are given to us with an immediacy, the force of which

is impossible to ignore. Finally, it is easy to note that ordinarily, our knowledge of our mental states does not follow the usual methods of investigation. We don't carry out an inquiry, lay out the evidence, weigh it against alternatives, ask other people's opinions on the matter etc. We only use our inner sense or self-scanning ability, which requires of us only that we focus on our current conscious experience and sometimes this may involve an effort of "looking harder". In summary, knowledge of phenomenal objects is sometimes given to us, but there are occasions when we have to discover it for ourselves by acting out as attentive agents.

One objection to this is that our perceptual abilities are made possible by our possession of the relevant apparatus, such as eyes and ears, whereas no similar apparatus seems to exist for our introspective capacity. If the absence of introspective capacity can be demonstrated by the lack of organs that perform the introspecting task, then we would also have to reject the ability to see dreams or to hallucinate, neither of which relies on our eyes and ears either. This is plainly absurd. A second line of attack insists that we can only have introspective ability if the distinct existences argument is accepted, according to which our introspective states are different from our beliefs about those states (Armstrong 1968, Shoemaker 1994). But since the two are not different, consequently the inner sense model must be false. I have already argued that our awareness of mental items is logically separable from judgments about said items. Recall how in the third section, I had demonstrated that we can be aware of the phenomenal character of a visual image and on the basis of this image, simultaneously have a false belief that the person we saw in the other end of the room is our friend. What separates the two is the logical precedence of the visual image over and above any judgements we might generate from them. Since the visual image can be real, but the judgment can be false, it follows that our introspective states are not the same as beliefs about them. From this I conclude that even if Shoemaker is correct in maintaining the distinct existences argument, the conclusion does not follow.

The last issue I want to tackle is whether there is any way in which first person authority can be salvaged after the findings that subjects can be sometimes unaware of their mental states and that they can possibly make mistakes about their mental states. The observational model of introspection seems to land an even more severe blow to the concept of first-person authority in the following way; if first-person authority is understood as the view that the subject has a more epistemically virtuous or secure (even if not infallible) position to gain acquaintance knowledge about the contents of their mind than the epistemic position occupied by a third-personal perspective about the outside world, then how can the perceptual model account for an epistemic difference between introspective judgments and ordinary perceptual ones? The idea is that since the observational model is the basis of both inner and outer perception, then there should be no difference whatsoever between the first person and third person perspectives, and the former cannot provide information that is qualitatively better or more secure than the latter. Consequently, we must either give up the observational model or give up first-personal authority.

I have one response and one rudimentary theory to offer. The response is that first person authority arises partially from the fact that it is exempted from at least one kind of error, namely that although subjects can form incorrect judgements about their mental states, they cannot be wrong that a certain mental state *feels like this*. That is to say that whereas knowledge of the external world is open to all manner of doubts, at least in principle, knowledge of the phenomenal character of demonstrative phenomenal items and their natures is not. This is what explains the epistemic security of first person-authority. The argument I want to offer in addition to the aforesaid response is going to be presented only in its outlines, and my purpose here is not to assert that the argument is necessarily correct, but only that it could prove why first-personal knowledge is more secure than the alternative. The argument is that first personal knowledge is obtained because the mechanism which contributes to it is sufficient to establish it. It is sufficient because knowledge of mental states is the result of a direct awareness of

the phenomenal character of mental items, and this is the only way in which such knowledge can be grounded. On the other hand, knowledge about the external world can go wrong even when the belief-forming mechanism behind is otherwise reliable. Since such a disparity exists between the two, we can conclude that the observational model of introspection is consistent with a version of first-person authority which affirms the impossibility of not knowing the phenomenal nature of mental contents, instead of judgments about them.

10. Conclusion

In this chapter, I have drawn the following conclusions. I have argued that knowledge of our mental items is a kind of acquaintance knowledge, since it fulfills the criteria for it. I have argued that knowledge of mentally private items should be understood as knowledge of the phenomenal quality of sensations and perceptions, rather than judgments about them. I have tried to argue for a moderate infallibility and indubitability, which concedes that although we can occasionally make mistakes about our mental states, we can still not doubt that a certain sensation *feels like this*. And finally, I have concluded that our direct access to our mental contents confers substantial but not unlimited first-person authority.

Chapter 3 – Final Considerations

Let us begin with a summary of what conclusions we have reached so far. We have found that i) mental privacy involves knowledge of one's mental states, ii) that such knowledge is possible, iii) that it is a kind of direct awareness of the phenomenal character of private items based on introspection, iv) that such knowledge may have certain limitations if it is construed in intentional terms but cannot fail to be infallible if it is a demonstrative containing terms such as "it feels like *this*", and v) that the first-personal authority it implies is based on privileged access an individual has to his own mental states which is a reliable mechanism of arriving at such knowledge of one's own conscious experience. In this chapter, I will focus on two things; first, I will consider whether it is plausible for two different individuals to have access to each other's mental items in one way or another, and thereby write off privacy as a philosophy myth and second, I will consider arguments based on what will be called "Burke's Assumption" to indicate what direction the future discussion of mental privacy might take.

11. Can somebody else know/have my thoughts and feelings?

If the privacy of the mental domain can be challenged, it can only be done so if other persons can have a way of knowing my mental states in a direct and non-inferential way. I have already suggested in the first chapter that observations of a person's behaviour will not do the task. There are two possibilities here in which this might work. The first would be if two or more persons have the numerically same mental experiences which coincide not only in their subjective natures but are also metaphysically one and the same. The second is to maintain that two or more persons do not in fact have the same metaphysically identical item in their stream of consciousness, but that there is still some other way in which one subject can transpose themselves to the other's consciousness such that while the

experiences continue to belong essentially to the first subject, the second subject can still experience it directly for himself and come to know it in the same way.

The first case would seem to require that two items or instantiations are meaningfully one and the same thing. It is not my purpose here to suggest any theory that offers a criterion of identity, but the absurdity of the proposal can be illustrated with simple examples. Imagine there are two chairs in the room, and both have the same qualities and attributes, such that it is difficult or impossible to tell them apart from each other. Surely, this cannot mean that there is only one chair present. It does not make much sense to insist that the identity of two items is dependent solely on the extent to which they share in the same pool of properties. If that were the case, then we would be compelled to uphold that two blocks of wood with the same dimensions, the same colours and mass, and perhaps even the same causal history (perhaps they come from the same manufacturer) have nothing to separate them. They really can't be the same because at the very least, they both cannot occupy the same space at the same time.

Conversely, if one were placed on a fixed point, and then removed and replaced by the second block at the same original spot, could we really say that it's the same block which we had placed before? I take it to be obvious that the answer must be no. I would suggest that the same applies to mental experiences since even if two phenomenal visual experiences are the same in every way (the same angles, the same vantage point of view), at the very least, we can meaningfully individuate the two by pointing out that the visual image belongs to two minds and two different bodies. From this, I conclude that the first proposal is not only logically impossible but positively absurd.

Arnold Zuboff and Peter Unger have proposed a thought experiment which could show that there is at least the metaphysical possibility that two persons can have quantitatively identical mental states (Unger 1990)¹. We are to imagine two people in different rooms who are molecule to molecule identical and have the same mental states and experiences. Suppose also that a zippering procedure is performed on both persons by which their brains are disembodied and the nervous system is replaced by radio

communicators, in such a way that although the brain is not in the body anymore, both twins continue to have the same phenomenal experiences as before. Imagine also that the brains are split into left and right halves and placed in three different locations; one for the body, and two vats for the halves of the brain. Both twins undergo the same procedure and both continue to have the exact same experiences. Now if a switch were flipped and the radio communicator made it so that one twin's left brain began to interact with the other's right brain and vice versa, and all this happens in a manner identical to how their left and right brain halves interacted with each other exactly as they did before the procedure started, then it seems as though the two twins end up having mental experiences that are not just qualitatively identical but numerically identical as well.

Putting aside the fantastical nature of this brain-split scenario, there are several reasons why it does not achieve the desired result. If there is to be complete numerical and qualitative identity, then it is necessary that the two twins must be identical in every way. However, when the procedure is performed, the two twins are still in different rooms and as such, are not numerically identical. Secondly, the case is liable to the same criticism that I alluded to earlier, namely that it is possible for one twin to survive if the other perishes. If one's body, left and right brain halves perished, and if the claim of numerical identity is correct, then how are we to account for the perishing of the one and survival of the other? We would have to conclude that nothing perished or survived, since both are supposed to be numerically the same, and yet it is clear that someone has indeed ceased to exist. Finally, it must be noted that we do not have very clear intuitions about what would happen if the thought experiment were to factually occur (Tye 1996). Since we don't have them, it seems inappropriate to use brain-split cases to draw any conclusions at all, whether for or against the possibility of metaphysical privacy. From these considerations, I conclude that we have no reasons to believe that it is logically possible for two individuals to have numerically identical experiences.

The second proposal has the advantage that it does not require a dubious criterion of identity, but its success depends on showing how one person can directly experience somebody else's conscious experience without having it himself. It is possible to imagine that there is some supernatural entity or oracle which gives me superpowers to enter into somebody else's consciousness in such a way that I can experience everything they do, without the ownership of those experiences passing on to me. Conversely, there could be a machine invented in the future, whose wires are plugged into two subjects so that one can "get inside" the other's mind and share his phenomenal experiences. My only response to this suggestion is that such scenarios are extremely unlikely and for all we know, it might be impossible for such a machine to be built. The most that this proposal shows is that it is at least metaphysically possible that one subject could witness the other's mind in a transparent way, but this can only happen if there is a way in which such a transposition could actually happen. If it cannot, then the proposal will collapse under its own weight. The burden of proof must lie on the claimant and it is impossible to assess it without any, but we can reasonably say that given the bizarre nature of these possibilities, the proposal does not seriously damage the mental privacy hypothesis for our present purposes.

12. "Burke's Assumption"

In this section, I will briefly respond to arguments that might appeal to what Edward Craig has called "Burke's Assumption" (BA) (Craig 1997). The assumption in question is that two distinct individuals may have the same internal states if their external states are identical. Two arguments can be made out of this. The first one would go along these lines; one could argue that there are natural phenomenal appearance properties that belong to objects in physical space, and since they are identical, we can use this to say that the mental states we enter into as a result of them will also be identical. But to say that physical objects may have natural properties which determine how they appear to us ignores the other

side of the perception process. Although things in the external world do not depend on us for their existence, mental items do. It would be senseless to talk of a pain that does not belong to anyone, or to assert that the phenomenal appearance of objects is not the product of an interaction between matter on the one hand, and the relative, contextual modalities associated with the perceiver's subjectivity on the other. In other words, mental items are dependent particulars and as such, how phenomenal appearances are determined by nature of the mind, which an individual has. Consequently, it would be mistake to talk of any natural phenomenal properties which belong to objects with spatio-temporal coordinates alone.

But there is a second way in which the argument might be construed; one that recognizes external similarities not as some natural kinds but physiological similarities between agents. You and I are both humans, both of us are the sum of the very similar physiological facts, and our bodies are similarly constructed. Given this, is it not reasonable to say that we have the same visual sense data when we look at the same object? The similarities in our external states (that is our bodies and organs) can reasonably lead us to having the same phenomenal items in our mind. I find this second proposal compelling but do not purport to defend it or lay out a plan of how it would work. Still, there are some concerns here which are better spelt out. If the argument is to have any persuasive force, then one would need to be committed to the view that physical similarities in our bodily constructions account for similar phenomenal output. How might this be proved? In the first chapter, the second formulation of epistemic privacy dictated that we cannot meaningfully compare similarities or dissimilarities in our conscious experience. Without this, how can we know that our similar external states *do in fact* lead to similar mental states? Additionally, even if there is such a relationship of entailment, how might we account for the fact that one subject might be necessarily prevented from knowing the phenomenal character of a particular kind of mental experience, and thereby fail to know what that experience is like for somebody else? Finally, would this line of reasoning not also be inferential in at least the strict sense

that it involves inferring the state of other subjects when they see a colour on the basis of what one state I have when I see it? In other words, it seems that I will still not be directly aware of someone else's mental state, nor would I have their mental state in a quantitative sense. All I could do (given that the assumption is reasonable, is that I could know what your state is without experiencing *your* particular experience. None of these objections are intended as final verdicts, but only to point out that an account of anti-privacy that appeals to BA still would face many challenges. Whether these challenges can in fact be met or not, remains to be investigated in future philosophical discourse.

13. Conclusion

In this chapter, I have tried to show that we do not have any strong arguments at the present moment which can undermine mental privacy. I have argued that other person's cannot have my conscious experience and if that is impossible, then the notion of epistemic privacy remains intact. I have also briefly considered two versions of Burke's Assumption, rejected one and pointed out concerns about the second without writing it off definitively.

Conclusion

I have attempted to do two things in this paper; first, to try to develop a coherent and detailed account of what mental privacy is supposed to be and second, to evaluate whether the concept is true. In the course of this paper, I have arrived at the following conclusions. Mental privacy devolves into two claims; epistemic and metaphysical. The kind of knowledge of mentally private items is non-inferential. Wittgensteinian and expressivist objections to the possibility of applying knowledge terms to our awareness of our mental contents are based on mistaken views about both the nature of mental privacy and the status of reports about first-person experiences. Mentally private items are best understood as elements in conscious experience which are directly given, and to which the subject has privileged access. Knowledge of private items is a form of acquaintance knowledge but this knowledge is necessarily non-intentional. Although we can typically make mistakes about judgments we form about what goes on in our minds, nevertheless we have a certain infallibility in knowing the phenomenal character of our feelings and perceptions, which remain fundamentally incommunicable. Our first-person authority with respect to our mental contents does not necessarily derive from a perfect ability to understand and report them, but rather out of the idea that we are in the best possible position to know their phenomenal nature, and this knowledge is formed out of a reliable process. Since a subject can know their minds in the way described, other people's ability to know them depends on whether it is possible for someone else to either have my mental experiences or somehow enter into my stream of consciousness but both of these are highly implausible.

These conclusions are by no means intended to provide a full account of mental privacy, which in any case is still a somewhat elusive concept. But I think I have sharpened the boundaries around the otherwise hazy edges of the concept to show that mental privacy is both plausible and coherent.

Bibliography

- Alston, W. (1971). "Varieties of Privileged Access." *American Philosophical Quarterly* 8: 223-241.
- Anscombe, G.E.M. (1981). "The First Person", in *The Collected Papers of G.E.M. Anscombe, vol. II: Metaphysics and the Philosophy of Mind*, Minneapolis: University of Minnesota Press.
- Armstrong, D. (1963). "Is Introspective Knowledge Incorrigible?" *The Philosophical Review* 72: 417-432.
- Armstrong, David M. (1968) *A Materialist Theory of Mind*, Abingdon, UK: Routledge.
- Armstrong, D. (1981). *The Nature of Mind and Other Essays*. Cornell University Press.
- Ayer, A. J. (1963). "Privacy". In *The Concept of a Person and Other Essays*. London: Macmillan.
- Baker, G.P. (1998). *The Private Language Argument*. *Language and Communication*, 18, 325–356.
- Bar-On, D. (2004). *Speaking My Mind: Expression and Self-Knowledge*, Oxford: Oxford University Press.
- Block, N. 1995. "On a Confusion About a Function of Consciousness." *Behavioral and Brain Sciences* 18: 227-247.
- Chisholm, R. 1981. *The First Person*. University of Minnesota Press.
- Churchland, P. M. (1988). *Matter and Consciousness*. The MIT Press.
- Craig, Edward John (1997). "Meaning and Privacy". In Bob Hale and Crispin Wright (Eds.), *A Companion to the Philosophy of Language*. Blackwell.
- Descartes, R. (1641). *Meditations on First Philosophy*. In *The Philosophical Writings of Descartes*, trans. J. Cottingham, R. Stoothoff and D. Murdoch. Cambridge University Press, 1985.
- Fales, Evan. (1996), *A Defense of the Given*, Lanham, MD: Rowman & Littlefield.
- Gallois, A. (1996). *The Mind Within, The World Without*. Cambridge University Press.
- Hacker, P.M.S. (1972). *Insight and Illusion: Wittgenstein on Philosophy and the Metaphysics of Experience*. Oxford: Clarendon Press.
- Hacker, P.M.S. (2005). "Of Knowledge and of Knowing that Someone is in Pain". In A. Pichler and S. Säätelä (Eds). *Wittgenstein: The Philosopher and his Works*. Frankfurt: Ontos
- Hacker, P.M.S. (1990/93). *Wittgenstein: Meaning and Mind. Volume 3 of an Analytical Commentary on the Philosophical Investigations. Part I. Essays*. Oxford: Blackwell.
- Lewis, C.I. (1946), *An Analysis of Knowledge and Valuation*, La Salle, IL: Open Court.
- Locke, J. (1689). *An Essay Concerning Human Understanding*. London: Thomas Basset.

McDowell, John (1982). *Criteria, Defeasibility, and Knowledge*. Proceedings of the British Academy, 68, 455–479.

McGrew, T. (1995). *The Foundations of Knowledge*, Lanham, MD: Littlefield Adams Books.

McGrew, T. (1999): “A Defense of Classical Foundationalism,” in L. Pojman (ed.), *The Theory of Knowledge: Classical and Contemporary Readings*, 2nd edition, Belmont, CA: Wadsworth.

Moser, Paul. (1989), *Knowledge and Evidence*, Cambridge: Cambridge University Press.

Raleigh, Thomas (2017). “Phenomenal Privacy, Similarity and Communicability”. *Ergo: An Open Access Journal of Philosophy* 4. PhilArchive copy v1: <https://philarchive.org/archive/RALPPSv1>

Shoemaker, S. (1963). *Self-Knowledge and Self-Identity*. Cornell University Press.

Shoemaker, S. (1990). “First Person Access.” In J. Tomerlin (ed), *Philosophical Perspectives 4: Action Theory and Philosophy of Mind* Ridgeview Publishing Company.

Shoemaker, S. (1994). “Self-Knowledge and ‘Inner-Sense’.” *Philosophy and Phenomenological Research* 54: 249-314. Reprinted in Shoemaker 1996.

Snowdon, P. (2011). “Private Experience and Sense Data”. In O. Kuusela and M. McGinn (Eds). *The Oxford Handbook of Wittgenstein*. Oxford: Oxford University Press.

Stout, Rowland (2010). “Seeing the Anger in Someone’s Face”. *Aristotelian Society Supplementary Volume* 84(1). <https://doi.org/10.1111/j.1467-8349.2010.00184.x>

Tye, M (1996). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. MIT.

Unger, P. 1990. *Identity, Consciousness, and Value*. Oxford: Oxford University Press.

Wittgenstein, L. (1953). *Philosophical Investigations (PI)*, 4th edition, 2009, P.M.S. Hacker and Joachim Schulte (eds. and trans.), Oxford: Wiley-Blackwell.

Wright, C. (1989). Wittgenstein's Later Philosophy of Mind: Sensation, Privacy, and Intention. *The Journal of Philosophy*, 86(11), 622-634. Retrieved June 22, 2021, from <http://www.jstor.org/stable/2027039>

Wright, C. (1998): Self-Knowledge: The Wittgensteinian Legacy. In C. Wright, B. C. Smith & C, Macdonald (Eds.), *Knowing Our Own Minds*. Oxford University Press.