



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

CREATIVE AGENCY

ANDREW JOSEPH MCKINLAY

Submitted for the degree of M Phil by research

The University of Edinburgh

2012

Declaration

I have read and understood the University of Edinburgh guidelines on plagiarism. This thesis was composed by me and is entirely my own work except where I indicate otherwise by use of quotes and references. No part of it has been submitted for any other degree or professional qualification.

Andrew McKinlay

Dedication

This work is dedicated to Monica Anne McKinlay, my mum, who in the summer of 1984, with care and imagination taught me to read and write.

Acknowledgements

I am indebted primarily to Tillmann Vierkant for his thoughtful and patient guidance and secondly to Dave Ward. Till and Dave's supervision has been invaluable in producing this work. Also deserving of thanks is the University of Edinburgh Philosophy department, especially Andy Clark and Duncan Pritchard, and the post graduate community who amongst many important influences Jonas Christensen , Connor McHugh, Anne Meylan, Andrew Hartline and Davide Fassio deserve a special mention. I also acknowledge the work and comments of my previous examiners. Many thanks also to the many proof readers who have given their time and efforts often without recompense: primarily Kay Barbour but also David Harris, Penny Thomas, Monica McKinlay, Jamie Black and Tom Roberts. The work was supported by a Warren McAlpine Scholarship.

Finally I would like to thank my family for the support they have given me. My mum and dad, Monica and John, and my brothers, Mike and Steve and their families. Their support both financial and emotional has been key, not only in producing this work, but throughout my life.

ABSTRACT

This thesis is about the agency involved with creativity. I am concerned particularly with the process of Insight Based Problem Solving (IBPS). IBPS is a problem solving process that is associated with a particular phenomenal experience, of sudden enlightenment, where new content arrives when the individual problem solver is not intentionally considering the problem at hand. The moment of insight is intuitively involuntary and as I argue for in chapter 1, is not successfully incorporated into any current account of creative agency. In chapter 2 I examine the scientific credibility of the process of IBPS. With such an account in hand I turn to the question of whether the process of IBPs is agentive. In chapter 3 I examine whether IBPS can be understood in terms of intentional action. In chapter 4 I examine whether we can be morally responsible for the products of IBPS based on previous mental actions. Neither of these approaches provides a satisfactory account of agency for IBPS because neither gives us a claim of moral responsibility for the specific products of IBPS. In chapter 5 I present a new positive account of the personal activity of IBPS. I argue that the moment of insight itself is a deliberative judgement that the content that is being brought to mind answers the question at hand. This new understanding of IBPS has implications for our understanding of mental agency, and the agentive nature of creativity.

Contents

Declaration:	i
Dedication:	ii
Acknowledgments:	iii
Abstract:	iv
Table of contents:	v-ix
Introduction: Creative Agency:	1-10
0.1 The Prologue	1
0.2 Definition of Creativity	1
-0.2.1 Creativity: Agentive or Non-agentive?	2
0.3 Agency	2
-0.3.1 Freedom and Responsibility	3
0.4 Insight Based Problem Solving	4
-0.4.1 Definition of Insight Based Problem Solving (IBPS)	5
-0.4.2 IBPS and its Importance for the Question of Creative Agency	5
-0.4.3 An Account of Agency for the Products of IBPS	6
0.5 The How Question	7
0.6 Thesis Summary	8
-0.6.1 Chapter 1: Creativity and Agency	8
-0.6.2 Chapter 2: The Science of Insight	8
-0.6.3 Chapter 3: The Bounds of Mental Action	9
-0.6.4 Chapter 4: Being Morally Responsible for the Products of IBPS	9
-0.6.5 Chapter 5: The Activity of Insight	10
0.7 Primer	10
Chapter 1: Creativity and Agency	11-37
1.0 Introduction	11
1.1 Creativity: a General Definition	12
-1.1.1 Intuitions about Creativity	13
-1.1.2 Creative Products	16
1.2 Creativity in Context	17
-1.2.1 Novelty Based on a Normative Judgment: Big C, little c, and Minimal c Creativity	18
-1.2.2 Conceptual Analysis: Valuable	19
1.3 A Necessary Agency Clause for Creativity	20

-1.3.1 Understanding Creativity as Broadly Intentional	21
1.4 Novelty and Psychological Creativity	22
-1.4.1 Boden’s Definition of Creativity	23
-1.4.2 Exploring and Transforming Conceptual Space	24
-1.4.3 Transformational Creativity, the Tripartite Model: Necessity and Sufficiency	25
1.5 Naturalising Creativity	27
-1.5.1 Can Transformation Separate Creative Cognition from Basic Cognition?	28
-1.5.2 The Psychology of Creativity: Six Ways to Study	29
-1.5.2.1 Psychometric Approaches	29
-1.5.2.2 Experimental and Biological Methods	30
-1.5.2.3 Computation Methods	31
-1.5.2.4 Biographical Methods	32
-1.5.2.5 Contextual Methods	32
-1.5.2.6 Conclusions on the Psychological Approaches to Creativity	33
1.6 Creative Problem Solving	33
-1.6.1 Where Does This Leave the Concept of Creativity?	34
-1.6.2 IBPS and Creativity	34
-1.6.3 The David Objection: What About an Agency Clause?	35
-1.6.4 Possible Dangers of Focusing on IBPS	36
Chapter 2: The Science of Insight	38-58
2.0 Introduction	38
2.1 An Intuitive Account of IBPS	38
2.2 Four Features of IBPS	39
2.3 The Special Process View is Essential for IBPS	40
-2.3.1 Evidence For and Against the Special Process View	41
-2.3.2 Early Evidences Supporting the Business as Usual View	41
-2.3.3 Cluster Analysis	44
-2.3.4 Feelings of Warmth	45
-2.3.5 Verbalisation Experiments	47
-2.3.6 The Light Bulb Priming Experiment	48
-2.3.7 Impaired Frontal Cortex Study	49
-2.3.8 Neuro-imaging, fMRI, EEG Studies of Insight	50
-2.3.8.1 Neurological Differences in Insight and Non-insight Problem Solving	50
-2.3.9 Conclusions on the Special Process Model	51
2.4 Is a Four Stage Model of IBPS Scientifically and Psychologically Credible?	51

-2.4.1 Preparation	51
-2.4.2 Incubation	53
-2.4.2.1 Evidence for Restructuring in Hindsight Bias Experiments	55
-2.4.3 Revelation	56
-2.4.4 Verification	57
-2.4.5 Constituents of Insight Based Problem Solving	59
 Chapter 3: Mental Agency and the Bounds of Action	 59-91
3.0 Introduction	59
3.1 Davidson: The Nature of Intentions	61
-3.1.1 Intentions as Reasons and Reasons as Causes	63
-3.1.2 Two Ways of Individuating Actions	64
3.2 The Simple View	66
-3.2.1 Strawson on Mental Action	68
-3.2.2 Is IBPS an Intentional Action Under the Simple View?	70
-3.2.3 The Access Puzzle	72
-3.2.4 The Consequences of the Access Puzzle for Understanding IBPS as a Simple Intentional Action	73
-3.2.5 Causal Connection within IBPS	74
-3.2.6 Conclusions: Mental Events and the Simple View	75
3.3 Advanced Causal Views	76
-3.3.1 Simple Problems? Searle's Answer: Two Types of Intentions	76
-3.3.2 The Nature of Intentions in Action	77
-3.3.3 Mele and Moser: The Advanced Causal Account	77
-3.3.4 How Intention-Embedded Plans Guide Intentional Action	78
-3.3.5 Mele's General Causal Theory and Mental Actions	79
-3.3.6 Strawson on Mental Action	81
-3.3.7 Memory Recall as an Intentional Action	82
3.4 Can IBPS be Understood as a Mental Action under GCT?	85
-3.4.1 Luck, Ballistic Parts and IBPS	87
3.5 Conclusions	90
 Chapter 4: Responsibility for the Consequences of Mental Actions	 92-109
4.0 Introduction	92
4.1 Moral Responsibility and Causal Responsibility	92

-4.1.1 Ascribing Moral Responsibility to Actions and Events	93
-4.1.2 Preparation as an Action	94
-4.1.3 Moral Responsibility for Action Consequences	95
4.2 Bratman: Intending to Act and Acting Intentionally	95
-4.2.1 Two Extensions by Means–End Reasoning	97
-4.2.2 Bratman’s View and Mental Agency	98
4.3 Responsibility for Foreseeable Consequences	101
4.4 Minimal Creative Thought: A Non-Accidental Agency Clause for Creativity	102
-4.4.1 Minimal Creativity and Responsibility for Creative Products	104
4.5 The Minimal Agency Model for IBPS	105
4.6 Moral Responsibility for the Specific Products of IBPS	106
4.7 Conclusion	108
Chapter 5: The Activity of Insight	110-137
5.0 Introduction	110
5.1 The Voluntary Nature of Mental States	111
-5.1.1 The Voluntary Nature of Judgements	112
5.2 Evaluative Control: Agentive Activity without the Paradigm Features of Agency	116
5.3 The Deliberative Stance	118
-5.3.1 Why is the Deliberative Stance Agentive?	120
-5.3.2 The Passive Nature of the Theoretical Stance Illuminates the Deliberative Stance	121
5.4 An Indirect Objection	123
-5.4.1 Managerial/Manipulative Control as Two Instances of Evaluative Control	125
5.5 The Story so Far	127
5.6 Deliberative Activity within IBPS	128
-5.6.1 Evaluative Control and IBPS	128
-5.6.2 The Deliberative Stance in IBPS	129
-5.6.3 Objection 1: Are the Processes Involved in Preparation Intentional?	130
-5.6.4 Incubation and the Activity of the Deliberative Stance	130
-5.6.5 The Deliberative Nature of Revelation	131
-5.6.6 Objection 2: Is There Really a Judgement within Revelation?	132
-5.6.7 Objection 3: Quick Judgements	133
5.7 Bringing it all Back Home: Conclusions on the Agentive Nature of IBPS	135
Conclusion: Concluding Creative Agency	138 -144
6.0 Preliminary Conclusion	138

6.1 Creativity and Agency	138
-6.1.1 The Relation between Creativity and the Agentive Nature of IBPS	138
6.2 The Science of Insight	140
6.3 The Bounds of Mental Action	141
6.4 Responsibility for Action Consequences	142
6.5 A New Way of Understanding Mental Agency	142
6.6 The Activity of Insight	143
6.7 Concluding Remarks	144
Bibliography	145
Appendix 1	152

Introduction: Creative Agency

0.1 The Prologue

Being creative, bringing about new ideas and novel products is a particularly important and rewarding aspect of our lives. Creativity is widely acknowledged to be essential for the arts (Gaut, 2010) from small personal affections, in the style of Bette Davis, to the fully fledged commitment to art we find in Georgia O’Keeffe, Bob Dylan, or James Joyce. But creativity is not restricted to the arts; it is central to how we sustain ourselves in adversity and progress both materialistically and psychologically. It plays an important role in the progress of science and mathematics, in the generation of technology and advancing philosophical understanding. In these and many more ways creativity is beneficial to society at large both in contributing to the advancement of all fields of human interest but also enlivening our everyday. Creativity is also personally rewarding, in the internal satisfaction that we derive from creative work and in the altered understanding of the world new ideas provide. But creativity, this important aspect of our lives, is often thought to be somehow estranged from us. It is a common thought that creativity is not something that we do but something that happens to us; that creativity results from the influence of outside forces, spiritual, or environmental, rather than from individual or group endeavour. Contrarily, at the heart of this thesis is the intuition that creativity, this intriguing cornerstone of our existence, is a central part of what it is to be an agent in the world. The goal of this thesis is to articulate this intuition. I will defend the claim that for novel and valuable products to be creative it will be sufficient that they are produced in a reason sensitive way. In establishing this claim I hope to make progress in dispelling the intuition that creativity is something beyond the endeavour of agents.

0.2 Definition of Creativity

In this thesis creativity will be used in two ways. It will be used in a general way to denote the topic. When it is used in this way I will refer to the *topic of creativity*. I will also utilize a technical definition of what it is for a product to be *creative*. In this sense, creative products will be any product that can be seen as both novel and useful and produced in a reason responsive way. In chapter 1 (sections 1.1 and 1.2, and subsection, 1.6.2 and 1.6.3) I present the case for understanding creativity in this way. I suggest the following interpretations of *novelty* and *value*. Ascribing novelty to a product will depend on that product being judged as novel within the context that it is produced. A minimal standard of novelty may be that a product has never occurred in that context before. The condition of value can be understood in terms of the interest products produce (see Boden, 1990, 1995). If products can be judged as interesting in the context they are generated then they have value. Finally, coming to an understanding of the way in which novel and valuable products can be understood as reason responsive is the main goal of this thesis. I will not present the case here. We can, then, consider individuals to display creativity or to be creative if they produce creative products.

0.21 Creativity: Agentive or Non-agentive?

Consider the following: an early caver pushes through a small gap in a limestone cavern in the Cambrian hills. Turning upright, he lights the torch he has carried for this occasion. Before him carved from the action of rainfall on the limestone, stands a statue of a man. Today we would recognize this statue as a facsimile of Michelangelo's David. Should this (imagined) feature of the natural landscape be considered creative? Most contemporary thinkers consider that for a product to be creative it must be novel and valuable. Under most definitions the limestone David statue will fulfil these criteria. Even so, most of us have the intuition that the limestone David should not be considered a creative product. The action of the rain on limestone doesn't seem to be the kind of thing that we consider can produce creative products. On the other hand we have no difficulty in considering Michelangelo creative or the statue of David as being a creative product. So, we might suggest that humans are the right kind of things; that they and they alone can produce creative products. This kind of thinking runs quickly into a type of prejudice unless we can explain why humans are the ones that can produce creative products. One answer to this question, one presented by Gaut (2010), Stokes (2011), and Klausen (2010) is that a novel and valuable product becomes creative by virtue of how it is produced, that part of the causal history of how a creative product is produced must be the activity of an agent. In the accounts of creativity of Gaut (2010), Stokes (2011) and Klausen (2010), being produced by an agentive process is seen as a necessary condition for products to be creative. Whether we want to understand agency as a necessary condition for creativity will depend on our intuitions about creativity (see section 1.3 and subsections 1.1.3 and 1.1.1). The particularly interesting question that arises from this line of thought is not whether agency is a necessary condition of creativity, but when and for what reasons should we consider creative products to be produced by agents. I will call this the question of *creative agency*. The question of creative agency asks by virtue of what are individuals agentive in producing creative products? This is an important question. In being able to understand creativity as personal to us we are able to associate creativity, the production of creative products, firmly with our own self concepts, with praise and guilt, with self esteem or the lack of it. If the creative process is agentive it will be a capacity of ours which, as I have suggested, is of great value. We can be considered beings capable not just of replication but as creators of our world, its environment and indeed the contents of our own minds. As such our understanding of creativity as agentive, or not, has important consequences for society and for individuals.

0.3 Agency

What does it mean to be an agent, or for a process to be agentive? One intuition is that *agency* is concerned with the behavioural events that we do, or bring about and these are distinct from those events that just happen to us. This distinction is close to the basic intuition about what actions are (see Wilson and Shephall, 2012). The question about what behaviour counts as an action is one way of addressing the question of how individuals are agentive. Actions, however, have taken on a life of their own. In

philosophy we have produced a large number of specific accounts of actions (see Mele and Moser, 1994; Bratman, 1984; Audi, 1993; and Ginet, 1990) and they cover a large area of agentic behaviour (I will discuss the relation of actions to a specific type of creative process in chapter 3). This type of behaviour has often been termed *voluntary* behaviour (Strawson, 2003) where voluntary behaviour is considered to be produced by the volition of an individual through an effort of will or an action. The term *involuntary* will be used in this thesis to denote any behaviour that is not voluntary as it is described above. But there are other instances of agency that are not actions in a philosophical technical sense. For example, we might think that the formation of beliefs is one example of a behaviour that is an agentic event even although it is not a standard case of intentional action (Hieronymi, 2009). We can however still make a distinction between agentic activities like these and those events that just happen to us and that we are passive towards (advocates of such accounts include Moran, 2001, Hieronymi, 2009, and Smith, 2005). So we have a range of behaviours that are involuntary in this sense yet still agentic. These behaviours can be distinguished from events, such as the experiencing of pain, that are passive or non agentic. We will explore this distinction and a particular kind of deliberative agency proposed by Moran (2001) in chapter 5. It is important for this thesis to place this type of deliberative activity within a philosophical context.

0.3.1 Freedom and Responsibility

The concepts of freedom and responsibility have been discussed in many guises¹. Topics that have received philosophical attention include political freedom, free will, agency and *moral responsibility*.² All are concerned with freedom and responsibility but in differing terminology and, in some cases, in genuinely different senses. In this thesis I will suggest that an individual can be held morally responsible for an event if it is appropriate to praise or blame them for that event (I will discuss this fully in chapter 4).

In the debate about free will *incompatibilists* (libertarians and sceptics) try to confirm or deny respectively, a type of *metaphysical freedom*. They are in pursuit of an explanation of how it is possible to have choice or to make up your mind in a way that given the same starting conditions could have turned out otherwise, that you can be the deciding factor, the loci of agency, in controlling the outcome of your actions in a way that breaks the chain of determinism³. In breaking the chain, actions can be seen to originate with an agent. We can call this type of freedom metaphysical freedom. I will not discuss it further.

Like the sceptics, *compatibilists*⁴ (those that hold that there is a sense of freedom that is compatible with determinism) think that metaphysical freedom is not possible, but unlike some sceptics they are not concerned with this outcome for they think a genuine explanation of freedom and choice and

¹ See Berlin 'Four Essays on Liberty' for an overview of many of these guises, (1969).

² See Strawson (1962) 'Freedom and Resentment'.

³ For an indeterministic account of free will, see, Kane (1998), Ginet (1970).

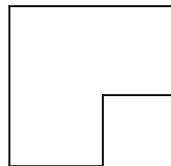
⁴ for some different compatibilist accounts, see, Frankfurt (1971), Dennett (1984), Fischer and Ravizza (1986 and 1998), Watson (2002), Wolf (1993) and Vierkant (2005).

agency is available even if we accept determinism. Compatibilists then, are searching for a different type of explanation to the question of how you can be free within a determined universe. One typical answer lies in accepting that freedom amounts to being able to bring about events in the world that can be understood as the product of or in virtue of an individual's reasons (for some discussion on the metaphysics of reasons see section 3.1.1). They accept a notion of responsibility and freedom that is understood in terms of an individual's reasons for doing what they do. In this thesis I will adopt this line. I will be concerned with agency in terms of how activities can be bound or controlled in some sense by an agent's reasons. I will not involve myself with other notions of freedom and in particular no agitation of metaphysical freedom will be undertaken.⁵ Our understanding of how and where reasons play a role in our activities as individuals will determine the extent to which we can consider creative processes agentive. Gaut (2012) suggests that creativity is agentive when the process can be understood as being sensitive to reasons. I will adopt this line, but will address the important question of how the phenomena of insight, of new content arriving suddenly from out the blue can be seen as being sensitive to an individual's reasons.

0.4 Insight Based Problem Solving

Insight based problem solving (IBPS) is the process of solving problems through insight, of being struck suddenly by new content that solves a problem seemingly without intentionally thinking about that particular question. (I will provide a technical definition below, Section, 0.4.1). The phenomenon is a common experience that sometimes occurs when considering difficult or unusual problems. Try to solve the figure problem by dividing the shape below into four parts that are the same shape and size. See appendix 1 (0.1) for the solution.

Fig.1



For many, when they engage with this task, they are initially unable to solve the problem. But after a moment's pause or inactivity they are often struck by a moment of insight where the solution, or an important aspect of the solution, suddenly occurs to them. This well known phenomenon is often called insight. It is associated with a particular phenomenal experience of revelation, of sudden enlightenment, that has been termed the '*Aha*' experience (Topolinski and Reber, 2010). This phenomenon of insightful problem solving has been linked with many important historical examples of creativity e.g. Archimedes' Eureka and Newton's apple. Further most of us will admit to having experienced this phenomenon and for

⁵ I do not discuss this topic because of the scope of the thesis. I do not produce an argument that libertarian freedom is un-associated with creativity. Kane (1985) suggests a connection between libertarian freedom and creativity but no formal account is provided.

some it is a regular occurrence. As such it has formed part of our understanding of what creativity is (Topolinski and Reber, 2010). IBPS can also be seen as one common way in which novel and valuable products are formed for the products of IBPS often provide a new understanding of the subject under consideration (subsection, 1.6.2). In focusing on IBPS, then, we are focusing on an important aspect of creativity (see section 1.6, for a discussion on this point). But I will follow Stokes (2007) in suggesting that instances of insight are not necessarily creative. Products produced by IBPS will be termed IBPS products. They may often be creative products, because IBPS often produces novel and valuable products but they are not necessarily so (subsection 1.6.2). To be so will depend on them meeting the independent definition of creativity i.e. being judged novel and valuable (see sections 0.2 and 1.6.1) and perhaps also on them being produced in an agentive way.

0.4.1 Definition of Insight Based Problem Solving (IBPS)

In this thesis I will utilise the terms ‘insight’ or ‘the moment of insight’ to refer to the specific episode of the ‘Aha’ experience within which new content is revealed to the problem solver. I will utilize the term insight based problem solving and the acronym IBPS to refer to a specific cognitive process with the following set of features (I argue for the scientific credibility of each of these features of IBPS in chapter 2).

- 1) IBPS solves problems utilising a different cognitive basis than non-IBPS (section 2.3).
- 2) Intentional processes within *preparation* (preparation is the first stage of IBPS, where we engage with the problem unsuccessfully) affects downstream problem solving (subsection 2.4.1).
- 3) Preparation can be distinguished from *incubation* (where incubation is a discrete stage of problem solving that utilises different processes from preparation and is not associated with any intentional activity (subsections 2.4.1 and 2.4.2).
- 4) Incubation takes place and has a positive effect on problem solving (subsection 2.4.2).
- 5) The ‘Aha’ experience takes place and is useful in indicating a specific type of problem solving (subsection 2.4.3).

From these features we arrive at the following definition of IBPS: a problem solving process that has three distinct stages. An initial stage of preparation that involves intentional activity; stage 2, incubation, a stage that utilises different cognitive processes than preparation and is devoid of an intentional activity and stage 3, revelation, where new content is realised in a moment of insight that is accompanied by an ‘Aha’ experience.

0.4.2 IBPS and its Importance for the Question of Creative Agency

In this thesis I call the particular question about the agency involved in creativity the question of creative agency. Gaut (2012), Stokes (2011) and Klausen (2010) attempt to answer the question of creative agency by addressing the concept of creativity as a whole. In Chapter 1 I will justify taking a different approach. I

suggest that we should consider how the question of creative agency can be addressed to a particular phenomenon associated with creativity, IBPS. My motivation for this approach is borne out of two concerns. First, the concept of creativity may be too broad to be associated with a single creative phenomenon. If creativity is associated with a distinct phenomenon, such as IBPS, then we will need to address how these individual processes are agentive. Second the intuitively involuntary and non-agentive nature of IBPS poses a distinct problem for accounts of creative agency, a problem that needs to be addressed directly with specific reference to IBPS.

IBPS is central to the question of creative agency because the moment of insight is one of the primary reasons why creative processes are often thought to be involuntary and non agentive. Products of insight just seem to come out the blue. If we accept the intuition that insight is involuntary and non-agentive then philosophers who think all creative process are agentive have a problem for insight has been strongly associated with creativity, further it often produces products that are novel and valuable. So we might think that we will either have to give up the intuition that creativity is always agentive or that products of insight should be considered creative. The latter seems particularly problematic for some instances of insight seem to be a prototypical example of creativity. In considering the former Gaut (2012) suggests that all creativity is rational because it can be considered reason sensitive. But he makes no claim of rationality for the moment of insight, rather considering the moment of insight non-rational. In order to provide an account of creativity agency we will need to address the issue that IBPS seems non agentive.

By focusing on IBPS we will be able to address this particular difficult aspect of the question of creative agency. We will also be able to extend the range of mental activities we can consider agentive (Section 5.7). This has implications for how we understand the activity of the mind and further helps us understand the psychology of the process of IBPS. As I have suggested the unique phenomenology associated with the moment of insight, the seemingly involuntary way content comes to mind, has provided much of the grounds for considering creativity as non agentive. IBPS is therefore a substantial target, for in addressing the agency question for IBPS we are addressing one of the aspects of creativity that is central to the intuition that creativity is non agentive. If this intuition can be seen to be false then we will have made significant progress in advancing the position that creativity (at least in these difficult cases) is a matter of agency. I will then, in this thesis, address the question of whether and how IBPS is agentive.

0.4.3 An Account of Agency for the Products of IBPS

Providing an account of agency for the activity associated with IBPS will not be easy. In chapter 3 I will examine whether IBPS can be understood in terms of causal accounts of intentional action. This will prove impossible. In chapter 4 I will examine whether we can be morally responsible for IBPS products because of intentional actions carried out in preparation. I will suggest that accounts on these lines are deficient in that they cannot provide a claim of moral responsibility for the specific products of IBPS.

In chapter 5 I will provide a new account of how episodes within IBPS that include the moment of revelation can be considered agentive. In order to provide this account I will look to Richard Moran's (2001) account of the deliberative activity associated with judgements. In Moran's (2001) account the personal activity associated with deliberation is not intentional but rather a particular type of first person activity that can be distinguished from the passive nature of the third person perspective (section 5.3 and 5.3.1). In this thesis I argue that certain episodes within IBPS are agentive not because IBPS is an action but because IBPS contains specific episodes of deliberative activity within preparation and revelation (section 5.6). Revelation, the moment of insight, itself is deliberative because it consists of a judgment that the new content that arrives through insight answers the question at hand. This account suggests that the moment of insight, where new content comes to mind, is the deliberative activity of an agent. This will provide us with a new understanding of the nature of this first person deliberative activity (section 5.7). Finally the nature of this activity will be discussed in the wider context of creativity and how it is possible for the products of IBPS to be both agentive and creative (section 6.1).

0.5 The How Question

A question of considerable importance for cognitive science is, how is it possible to come up with novel and useful ideas? What are the cognitive processes that produce them, how can we think things that we have never thought of before (Boden, 1990)? No good answer to this question is currently available. Many possible systems have been proposed. Many consider that cognitive restructuring and the making of new connections between old concepts is crucial for producing creative products (Ohlsson, 1984). As Stokes (2007) suggests such connections may be made unconsciously through the spreading activation of cell assemblages when we are attending to other things. In section 1.4 I explore several mechanisms proposed by Boden that are important for her attempt to define the concept of creativity as a mechanistic process. Further in chapter 2 (subsection 2.4.2.1) I present some evidence that restructuring is important for IBPS. These accounts are included in the former case because it is a necessary component of Boden's definition of creativity and in the latter to support the case for understanding IBPS as a discrete cognitive process. No direct attempt is being made, at any time in this thesis, to answer the question of the cognitive nature of how new answers are generated. The question that will concern us is the question of creative agency, as it is directed at IBPS. It might be suggested that to answer the question of creative agency we need to understand the how question. I will suggest that since we have independent scientific support for the causal interaction of the stages involved in IBPS we can address the question of creative agency without understanding the exact details of how the how question might be answered.

0.6 Thesis Summary

For the convenience of the reader a summary of each chapter and a description of the role each chapter is playing in the thesis are presented below.

0.6.1 Chapter 1: Creativity and Agency

Role in thesis: Chapter 1 provides a justification for examining the agency question in relation to IBPS rather than the concept of creativity as a whole.

Summary: In chapter 1 I introduce the concept of creativity. I suggest that creativity conceived as the production of products that are both novel and valuable is too widely applicable to pick out a single creative phenomenon. I discuss the conditions of novelty and value as they have been pursued in the philosophical and scientific literature and suggest that none of these approaches successfully pick out a singular or group of cognitive processes that can be seen as creative. I examine the attempts to provide an account of agency as a necessary condition. Gaut's (2012) account suggests that the way forward in providing an account of creative agency is to understand how creative phenomenon are reason responsive i.e. directed and controlled by an individual's reasons.

One process that is problematic for an account of creative agency based on reasons responsiveness is IBPS. IBPS is intuitively associated with creativity but it is also intuitively involuntary and non agentive. This poses a problem for Stokes (2011), Gaut (2010) and Klausen (2010) who suggest that agency is a necessary feature of creativity. One way of responding to this difficulty is to directly address whether IBPS is agentive. IBPS is historically connected to the concept of creativity because the unique phenomenology of ideas, seemingly striking out of nowhere, has been included in many important accounts of creative inspiration. IBPS, then, is a prototypical example of a creative phenomenon that seems involuntary and non-agentive. So to address the question of creative agency we need to address the phenomenon of IBPS. We need to ask the question: in virtue of what is the process of IBPS agentive? In providing an answer we can go a long way to answering the objection that creativity is non-agentive.

0.6.2 Chapter 2: The Science of Insight

Role in this thesis: I suggest that IBPS is a phenomenon that is associated with creativity but is intuitively involuntary and non agentive. This poses a problem for philosophers who think creativity is agentive. One possible response to this difficulty is to deny that IBPS is a distinct process. In chapter 2 I examine the scientific case for understanding IBPS as a discrete cognitive process. I confirm that IBPS should be seen as a discrete cognitive process. Therefore any investigation of creative agency will have to address the question of whether IBPS is agentive.

Summary: In chapter 2 I set out the scientific case for understanding IBPS as a discrete cognitive process. Within the scientific literature there are two standard positions on the nature of IBPS. These are the special process view and the business- as-usual view. The business-as-usual view suggests that IBPS is just like other problem solving, that might on occasion be accompanied by the weird phenomenology of revelation. In contrast, the special process view suggests that IBPS is involved with a distinct cognitive function that resolves problems in a distinct way to non-IBPS. I set out the early case against IBPS (Weisberg and Alba, 1981) and provide more recent evidence that suggests the Weisberg and Alba (1981) case against the

special process view is, at best, inconclusive. I then present the six main areas of evidence for and against the special process view: feelings of warmth (section 2.3.4), cluster analysis (2.3.5), impaired frontal cortex studies (2.3.6), verbalisation studies (2.3.7), priming studies (2.3.8) and neural imaging data (2.3.9). Overall the data supports the conclusion that IBPS can be understood as a singular cognitive mechanism. Supporters of creative agency will therefore have to address the question of how IBPS is agentic. With a basic view of IBPS supported by the scientific evidence in hand we have the appropriate foundation to begin to examine that very question.

0.6.3 Chapter 3: The Bounds of Mental Action

Role in thesis: The role of chapter 3 is to explore the extent to which the process of IBPS can be understood as agentic according to traditional accounts of intentional action. With no convincing account forthcoming the chapter legitimizes looking further afield for an account of the agency involved with IBPS.

Summary: I begin chapter 3 by presenting Galen Strawson's (2003) account of mental action. In Strawson's (2003) account several types of mental activity, such as priming and shepherding actions will be intentional. I then examine a less stringent interpretation of the simple view. Here we will encounter some of the major difficulties of understanding mental events in terms of intentional action. One of these difficulties is the access puzzle. The access puzzle occurs when the content of an intention cannot be linked with the result of the mental activity because the agent does not know what the result of the mental activity will be, and so cannot intend to bring it about (Proust, 2001). This is problematic for IBPS because the goal of IBPS is to bring about new content that answers the problem at hand. I then examine more advanced causal accounts of intentional action, particularly Mele and Moser's (1994) account and Mele's (2009) accounts but because of the access problem none of these accounts can provide a satisfactory account of intentional action for IBPS.

0.6.4 Chapter 4: Being morally Responsible for the Products of IBPS

Role in thesis: Chapter 4 examines whether an account of moral responsibility for IBPS products is possible if we consider preparation as containing intentional actions. As such it completes the investigation of the extent that an account of agency for IBPS can be understood as the consequence of intentional action.

Summary: I introduce the concept of moral responsibility. I examine Bratman's (1984) account of intentional action and accounts of moral responsibility for action consequences. I examine whether this type of account can provide an account of moral responsibility for the products of IBPS including Stokes' (2011) account of minimal creativity. I conclude that accounts on these lines are unsatisfactory because they do not provide a claim of moral responsibility for the specific content produced in IBPS.

0.6.5 Chapter 5: The Activity of Insight

Role in thesis: Chapter 5 provides a positive account of the agency associated with IBPS: as such it accomplishes one major goal of the thesis. With an account of IBPS in hand we can turn to examining how the agency within IBPS is associated with creativity more generally.

Summary: In chapter 5 I provide the positive case for understanding IBPS as agentic. I begin by discussing the voluntary nature of intentions and judgments. I understand judgments to be a matter of committing to particular attitudes. I agree with Hieronymi (2008) that judgments cannot be intentional actions because they are restricted to a set of reasons that are answerable to the truth. In order to understand the agentic nature of judgements I adopt Moran's (2001) account of active deliberation where the activity of the agent is associated with coming to decide transparently on what you will believe. I argue that IBPS contains two episodes of deliberative activity within preparation and within revelation. Within revelation the moment of insight itself should be seen as deliberative judgment that the new content is an answer to the question under consideration. In coming to view revelation as an episode of deliberative activity I argue that our active participation in IBPS cannot be restricted to different forms of intentional action but includes episodes of deliberative activity that contain no intentional element. The agentic nature of IBPS can be usefully grounded in the particular deliberative activity of coming to decide for yourself to commit to particular content as providing an answer to the question the individual engaged in IBPS is pursuing. In the conclusion I provide an overview of the agentic nature of IBPS. I then turn to examine the wider implications that an understanding of IBPS as a deliberative process has for creativity as a whole.

0.7 Primer

In the first chapter we will turn to the concept of creativity. I will suggest that in order to examine the nature of the agency involved in creativity we will achieve most by focusing on the particular phenomenon of IBPS. IBPS will then be the focus of our attention for most of the thesis. Chapter 2 will establish its scientific credibility, while in chapter 3 and 4 I will examine the agency involved with IBPS by considering whether it can be understood as an intentional action. Finally in chapter 5 a positive account of the agentic nature of specific episodes within IBPS will be presented that links the moment of insight with a deliberative judgment. If you are ready to begin P.T.O.

Chapter 1

Creativity and Agency

1.0 Introduction

This chapter aims firstly to introduce the topic of creativity in terms of both the philosophical and psychological traditions and secondly to provide a justification for the approach I adopt in this thesis to explore the association between creativity and agency. Within the philosophical literature it is widely accepted that for a product to be creative it must qualify under some standard, as both novel or original⁶ and valuable (Gaut, 2010). Understanding creativity in this way motivates a range of definitions of creativity (for some authors who utilize this definition see Amabile, 1983; Boden, 1990; Gruber and Wallace, 1999; Lurnsden, 1999; Martindale, 1999; Mumford and Gustafson, 1988; Pfeiffer, 1979; Sternberg and Lubart, 1999; Unsworth, 2001). Others (Stokes, 2011; Gaut, 2010; Klauson, 2010) consider that novelty and value alone are not enough. They suggest that creativity must be in some sense an agentive phenomenon, one that can be understood as the activity of an agent. Here, agents can be understood as individuals who are personally associated with their acts and activities where it can be said that those acts or activities are something that the agent does rather than something that happens to them. One question that concerns these authors is whether agency is required for creativity. But perhaps the more important question is what I have called the question of creative agency: in what way is creativity agentive? Why should we consider the production of creative products to be the activity of an agent i.e. sensitive to or controlled by their reasons? In considering the question of creative agency we need to come to an understanding of what the concept creativity refers to. I will suggest that creativity should be considered a widely attributable concept that depends on an individual's thoughts or products being judged as novel and useful by individuals or communities under some normative standard⁷. If this is the case, it will be unlikely that creativity can be associated with a single phenomenon. This will prove problematic for the question of creative agency because distinct phenomena associated with creativity may demand individual accounts of agency.

We can loosely call phenomena that can generate novel and useful products *creative phenomena*. Examples of such phenomena include problem solving, both IBPS and Non-IBPS (NIBPS). I will emphasise the importance of understanding the creative phenomenon of IBPS. There is a wide range of creative phenomena (methods of producing creative products) such as free sketching⁸ or the ability to produce novel and useful products through painting or dance. Many of these creative phenomena may rely on distinct cognitive mechanisms or distinct modelling of the interaction between mechanisms that will be

⁶ I will use novel rather than original for it has been utilised widely in both the philosophy and psychology of creativity.

⁷ For the moment we can leave aside a further condition of agency that might be required to consider novel and valuable products as creative.

⁸ An example that Lubart (1999-2000) has suggested does not fit neatly into a problem solving model.

important to understand if we are to make an appropriate ascription of agency; that is, if we are to correctly understand how these processes are controlled by an individual's reasons.

We might think that in order to distinguish different creative phenomena we need to associate them with distinct cognitive processes. If this cannot be done then we may have only superficial reasons for thinking them distinct. The aim of chapter 2 is to establish the distinct cognitive basis for IBPS. As such, it can be seen as a distinct creative phenomenon to NIBPS. I will suggest that if we have a distinct range of creative phenomena the best approach to investigating the relation between creativity and agency is to give up on a single account of how all novel and valuable products can be brought about in an agentic way. Instead philosophers interested in the relation between creativity and agency should focus on understanding the agentic nature of particular creative phenomena. Following my own advice I will, throughout this thesis, concentrate on understanding how the creative phenomenon of IBPS is agentic.

IBPS is a good candidate for such an approach because it can be understood as a particular cognitive process that is scientifically respectable (Chapter 2 will be dedicated to supporting this claim) and has a central place in our thinking about creativity. Moreover, for our specific interest of creativity and agency, IBPS is a particularly problematic and interesting case, for it seems that insights arrive out of the blue when we are not actively in pursuit of a problem. The heart of IBPS, the moment of insight, is intuitively involuntary and non agentic. We therefore have a prima facie reason to doubt whether an ascription of agency based on intentional action is appropriate to this process. (In chapter 3 I will spend a great deal of time, establishing the extent of this claim). Providing an account of agency for IBPS with its peculiarities of engagement and access is a substantial obstacle for any philosopher who thinks creativity is agentic. I will tackle this obstacle in chapter 5 where an account of how specific episodes of IBPS, including the moment of insight itself are deliberative. Establishing such an account will clear one of the main hurdles in seeing creativity generally as agentic.

1.1 Creativity: A General Definition

In introducing creativity I will begin with a brief look at the history of the word and its usage⁹. The word *create* is of Latin origin (Pope, 2005). It originally denoted the act by God of bringing about the heavens and the earth, *ex nihilo*, from nothing (Pope, 2005). The original creator then, in this sense was God and he brought into being all of creation. So the term creation can stand for all the heavens and the earth, and the word 'creature', although more often associated with animals, can denote any part of that creation (Pope, 2005). Since these Latin origins the phrase 'to create' has taken on the more general meaning of making something new. 'New' here is given a restricted understanding of being original or novel. An individual who makes a new wheelbarrow can be said to be its creator but she will not necessarily be

⁹ For a full history of the words create, creative, creativity, creature and other related words see Pope (2005); for a brief entomology see Götz (1981); for an alternative history focused on creative research see Mayer (1999).

called creative. To be creative an individual will have to bring about something novel or original. Creativity in this context has commonly been associated not just with the production of new objects but also new thoughts. It is this sense of creativity, the bringing about of new or original thoughts or objects that has been at the core of the investigation by philosophers and scientists. In this context artists have been strongly associated with the term, and this might be because their works are often thought to be more original than the work of other fields. We consider artists, whether painters, musicians, dancers etc. to be creative people, to be part of the creative industries, and to produce work that is creative. But we seem also to identify creativity in the work of scientists, of mathematicians, or (dare I say) philosophers and a wide variety of work and professions that extends to the working and playing practices of a wide cross section of the populace. Creativity then is the property of being creative, bringing about originality or novelty, in whatever field or industry.

1.1.1 Intuitions about Creativity

When we first consider what the concept of creativity means, we are motivated by its linguistic application where, as we have discussed, creativity seems broadly applicable to a wide range of endeavours that bring about new thoughts, actions or objects. It is also often associated with a particular type of experience, a moment of realisation where ideas strike us, seemingly from nowhere. This final type of phenomenon we will call the moment of insight and it is one of the hallmarks of IBPS. Thinking and acting creatively, understood in this broad sense, is part and parcel of how people go about, and make up their lives. Yet creativity if widespread does not seem to be commonplace. The work and actions produced by creative people, in whatever field, are not what we would call ordinary; they stand out, interest, and even surprise us. If creative events and products have anything in common, it is their originality or novelty. They contain some new, novel element. This primary feature of creativity is what separates creative events and products from the whole other bundle of everyday happenings, be it putting out the light, the garbage or someone's eye that we normally attribute to agents. We can then suggest that what separates creative acts or products from everyday products or actions is their novelty; that they are, in some sense or context, new. Here 'novelty' will be used instead of 'new' or 'original' because that has been the standard word used in the literature (see Mayer, 1999 for a table). I will utilise it to try to capture the difference between everyday and creative events. We will soon look at the different ways some philosophers have tried to establish this distinction (sections 1.2 and 1.4).

It might seem that novelty and creativity are synonyms, but we may have reason to think this is not so. Consider the example, inspired by Kant, of a monkey, typing 200 words in a random order. Intuitively most of us have difficulty in ascribing the term creative to this monkey's work. There are two different ways the example may fail our intuitions.

Firstly it seems that the product the monkey produces is of no value. Value can be cashed out in different ways (Schroeder, 2012); many consider that creative ideas should lead somehow to some form of

extrinsic value, that they should be valuable for reaching other ends (Gaut, 2010). Some consider that creativity is a virtue and so has ethical or epistemic value in being a kind of virtue¹⁰. The ascription of different values towards acts of creativity and creative products is not necessarily seen as a condition for a novel act to be a creative one (see subsection 1.2.3 for a discussion of value as it relates to creativity). But most philosophers (Gaut, 2010) have supported the need for some response to the problem that intuitively not all novel products are generally considered creative¹¹. This has most commonly been presented as a necessary condition. So on an understanding that includes value and novelty, creative products would have to be novel, and they must also be valuable, in some sense. Novelty alone will not suffice, nor will value, for many everyday actions have value without us considering them creative, e.g. barring any unusual circumstances, putting out the cat. In providing a definition of creativity I will stick with the generally accepted intuition that creative products (will at least) need to be novel and valuable.

But there is a second sense in which the above example fails to satisfy our basic intuitions about creativity. Consider a modified example. A monkey is at a typewriter, tapping away, you rip out the paper to find: 'A way a lone a last a loved a long the riverrun, past Eve and Adam's from swerve of shore to bend of bay, brings us by commodious viscous of recirculation back to Howth castle and Environs...'. Checking the pile of paper beside the typewriter you find the complete transcript of Joyce's (1940) *Finnegan's Wake*. This work typed out at random before Joyce wrote it, forbidding any doubt about its brilliance, is novel and valuable. Yet to describe it as creative when it arrives from such a random occurrence (the random typing of a monkey) seems to miss part of what we consider important about the creative nature of the work. Here the failure to satisfy our intuitions seems to stem, not from any doubts about value, but from a failure of agency. It is a failure to be able to ascribe the creative product to the activity of a rational agent.

The thought that the generation of art requires agency and rationality is found in the work of Kant (Gaut, 2010; 2012)¹². On the creation of Art (that can be understood as the bringing about of beauty), he suggests

'Abundance and originality of Ideas are less necessary to beauty than the accordance of the Imagination in its freedom with the conformity to law of the Understanding. For all the abundance of the former produces in lawless freedom nothing but nonsense; on the other hand, the Judgement is the faculty by which it is adjusted to the Understanding' (Kant, 2010 pp.120-121).

¹⁰ See Audi (2001) also Roberts and Wood (2007) and Zagzebeski (1996).

¹¹ See Stokes (2008) for a different condition 'An *F* is creative only if *F* could not, relative to the cognitive profile of the agent in question, have been done or performed (by *A*) before the time it actually was'.

¹² Stokes (2008) suggests that in certain contexts according to Kant the work of a genius should not be considered rational.

For Kant reason is important in bringing about beauty. Here rationality can be seen as the basis of agency. Events that cannot be attributed reasons are not attributable to agents¹³. This inclusion of rationality in art is in contradiction to a separate line of thought on artistic creation, first expressed by Plato.

Socrates: For all good poets, epic as well as lyric, compose their beautiful poems not by art, but because they are inspired and possessed. And as the Corybantian revellers when they dance are not in their right mind, so the lyric poets are not in their right mind when they are composing their beautiful strains: but when falling under the power of music and metre they are inspired and possessed; like Bacchic maidens who draw milk and honey from the rivers when they are under the influence of Dionysus but not when they are in their right mind. And the soul of the lyric poet does the same, as they themselves say; for they tell us that they bring songs from honeyed fountains, culling them out of the gardens and dells of the Muses; they, like the bees, winging their way from flower to flower. And this is true. For the poet is a light and winged and holy thing, and there is no invention in him until he has been inspired and is out of his senses, and the mind is no longer in him: when he has not attained to this state, he is powerless and is unable to utter his oracles....

Ion. Yes, indeed, Socrates, I feel that you are right; for your words touch my soul, and I am persuaded that good poets by a divine inspiration interpret the things of the Gods to us (Plato, 2004 pp.8-9).

Here it is madness or irrationality not reason that produces creative products. This line of thought has been prevalent¹⁴, sometimes counter productively so, throughout the philosophical discourse on creativity. These two intuitions, that creativity comes from madness and that creativity is attributable to a rational agent, are in conflict, a conflict that has been commented on recently by Berys Gaut (2010, 2012).

For Plato, the Poet is a vessel for the gods; as such the poets themselves are not the creators. Their work originates from the gods and the poet is the vessel through which the work is brought into the world. The poet's reason plays no role. This poses a difficult question for the modern philosopher: if creative acts or happenings are not based on reason in what way can they be agentic? This question may be particularly problematic for the process of IBPS, where the moment of insight seems to come when we are not engaged in thinking about the problem at hand.

Those convinced by Plato may well feel that this suggests that there is no creative agency. But this may be hasty. Throughout this thesis I will explore the question: how is it that we can be agentic in regard to IBPS? We will investigate different ways IBPS can be connected to an agent's reasons, by trying

¹³ For other accounts of rational creativity see Aristotle (1984) Jon Elster (2000).

¹⁴ See Schopenhauer (1969), Nietzsche (1967) and more recently Levinson (2003).

to understand it as a form of intentional action (chapter 3) or a consequences of intentional action (chapter 4) and finally coming to understand part of the process of IBPS, the moment of insight, as a deliberative reason sensitive activity (chapter 5). After we have addressed the question of the agentic nature of IBPS we will consider how IBPS products relate to creativity. But before we can investigate the nature of agency involved with IBPS and its association with creativity we need to get clearer about what is meant, if anything, by creativity.

1.1.2 Creative Products

Within this thesis, and this chapter in particular, I will talk about *creative products*. Here creative products will not just stand for the physical objects generated by the creative process, such as paintings or computers but also for the thoughts and mental states on which such objects are based. There is some debate about what should be the bearer of creativity (Klausen, 2010; Götz, 1981; Stokes, 2008) and whether we should understand the property of creativity to be attributable to products, processes or people or perhaps even groups or societies (Paulus and Nijstad, 2003). Overall creative products have been widely considered the easiest target to assess (Klausen, 2010). This may be the case because it is difficult to separate ‘creative’ processes from the products they produce (Amabile, 1996). But there has also been some objection to only talking about creative products in that this may disqualify creative thinking that does not lead to the production of a product¹⁵ and that talk of products is too focused on a western understanding of creativity¹⁶. Stokes (2008) suggests that assessing creative products alone without considering how they are produced will not provide a good assessment of their creativity; creativity although always associated with producing a product can only be properly assessed by considering the process through which the product is produced. I do not want to get into this debate. I will talk about creative products. I will consider creative products widely, to include mental states. In so doing many of the perceived problems of understanding creativity as associated with products can be avoided. Further, in talking about creative products I am not committed to the view that we can understand creativity solely as the physical nature of the products. We can consider that creative products are creative by virtue of the way they are produced as well as their relative novelty and value. Indeed ascribing a condition of agency to creative products will demand that we understand how those products come about.

As a matter of clarification within this thesis creative products are to be understood separately from *IBPS products*. IBPS products are the products of the process of IBPS. IBPS products (depending on the definition of creativity) will often be creative but are not necessarily so (I defend this claim in subsection, 1.6.2). In the conclusion of the thesis after much work has been done understanding the

¹⁵ For an account of creativity that does not demand the production of products see Bailin (1984).

¹⁶ Eastern conceptions of creativity have historically focused on the process not the product (Lubert, 1999).

process and agentic nature of IBPS, the relation between IBPS products and creative products will be discussed.

1.2 Creativity in Context

So far in discussing the concept of creativity I have highlighted the three criteria of novelty, value and agency. I have suggested that each might be seen as a necessary condition of creativity. In considering the conditions of novelty it becomes apparent that the term needs to be contextualized. What is novel in one context is not necessarily novel in another. Value too might need to be contextualised depending on the type of value we have in mind. But I will for the moment concentrate on novelty because the context in which creative products are produced has been used as a way of assessing novelty.

Margret Boden (1990) provides a distinction between *Historical (H-) creativity* and *Psychological (P-) creativity* to try to provide a way of capturing two important but distinct contexts for novelty. A product is H-creative if it has never been brought about before. For instance, Albert Einstein's theories of relativity can be seen as an example of H-creativity because in the history of the human species no one had thought of this theory before. The context for judging creative products as H-novel, and therefore possibly H-creative, is that they occur for the first time in human history. In Psychological creativity (P-creativity) the context is the psychological process of the individual. For a product to be an example of *P-novelty*, and therefore a possible candidate for being P-creative, what is important is whether or not the product an individual produces is new to them; that the creative product has not or perhaps could not have been produced by them before. Accounts of both P and H creativity exist¹⁷.

A judgment concerning H-novelty will depend on historic facts. Thus, providing an account of how such ideas come about will be bound up with the historical account of ideas. While accounts of H-creativity may be useful in their own right they suffer against the prevailing tide of naturalism¹⁸. For a philosopher interested in the mind a historical explanation alone will not be sufficient, for the explanation excludes a general understanding of how creativity relates to most individuals¹⁹. I will therefore not pursue an account of H-creativity.

¹⁷ For an account of P-creativity see Boden (1990), for an account of H-creativity see Feldman et al. (1994).

¹⁸ Boden herself considers that providing a good account of H-novelty will ultimately be flawed for the following three reasons: firstly many non-psychological facts are relevant and some cannot be discovered, secondly these facts are not all of the same type, so it will be hard to provide a unifying theory, and thirdly the values that might be associated with novel ideas differs across cultures and societies; they are inconstant, and unpredictable, even within social groups (Boden, 1990, 1995).

¹⁹ Boden provides the example of Alan Turing's discovery of a mathematical theory that he utilised in his application for Cambridge. The work had recently, independently, been discovered by an established Scandinavian mathematician. We are left wondering what it is about these events that we want to explain. An explanation of why the established mathematician is H-novel seems not to be associated with the phenomenon by which he produces the new breakthrough. It seems to Boden (1990) that an account of H-creativity for these endeavours leaves much still to be explained about both cases.

1.2.1 Novelty Based on a Normative Judgement: Big C, Little c, and Minimal c Creativity

One major disadvantage of only focusing on H-creativity is that it focuses our attention on a very small sample of extreme cases. If we consider creativity to be a common if non-typical event, then in only considering creative episodes that generate H-creative ideas we are disregarding a large number of intuitively creative happenings.

One way of alleviating the difficulty of H-creativity being too selective is to introduce more creative categories. One common distinction is between Big *C-creativity* and little *c-creativity* (Kozbelt, 2010). C-creativity is similar to H-creativity but is defined in terms of creative products (Beghetto and Kaufman, 2007). So a creative product is C-creative if it occurs for the first time in human history and satisfies a particular value condition. On the other hand for a product to be c-creative it is required to be judged in the context at which it occurs as novel and it also needs to be valuable. Finally *Minimal c-creativity* moves the focus away from creative products onto the creative process. For an event to be minimal c-creative it must be novel and valuable to the individual who produces it. Examples of minimal c-creativity may include specific episodes of concept learning (Beghetto and Kaufman, 2007). Minimal c-creativity then is similar to P-creativity in that it moves the context of how novelty is to be assessed to the psychological process of the individual. I will examine the success of this move to the psychological context below (subsections, 1.4 and 1.5).

First, will an account of creativity in terms of C and c-creativity be helpful in directly addressing the question of creative agency? In both cases assessing whether it is C or c-creativity will require a judgement to be made as to whether the product is sufficiently novel to be considered creative. Ascribing creativity based on a judgement of whether a product is novel or not will not help us with the problem that unique creative phenomena may require distinct accounts of agency. For such a definition of creativity cannot be associated with a single creative phenomenon. An individual who arrives at a product by trial and error will be as creative as an individual who solves a problem through insight. The product is being judged by its social (c) or historical (C) context, not by the way in which it is produced. Such accounts of creativity then will not necessarily be associated with any particular creative phenomenon that can generate novel and useful products. This may prove problematic if we wish to provide an account of why all the products judged as creative, on the basis of novelty and value, will also be agentive. Singular accounts of creative agency, that is explanations of why all creative happenings are agentive, will have to be associated with all the different phenomena that produce any product that is judged, in its particular context, as novel and valuable. Such a broad account of creative agency may be difficult. I will examine two attempts to provide this type of account after a brief discussion of value.

1.2.2 Conceptual Analysis: Valuable

Why should we consider that creative products need to be valuable? Consider the example of an inventor that creates what we can call a gizmo. The gizmo has many complex buttons and levers; in this sense it has never been invented before. But the gizmo has no practical application beyond being able to be utilised as a large heavy cumbersome club. The novelty contained in the gizmo, its weird combination of buttons and levers, has no value in the sense that it is not useful.

Most people have the intuition that such a construct is not a creative product. If we agree with this intuition then we might think that for a product to be creative its novel features must be of some value. Value is itself a very broad term. A distinction between *intrinsic value*, value that an object holds in its own right, and *extrinsic value*, value that an object holds in virtue of some other worth or value is prominent in the literature (Schroeder, 2012). Intuitively extrinsic value may be most commonly associated with creative products. Creative products should be useful to us. Usefulness itself can be defined in different ways. It may just reduce to the fact that a product provides some value to an individual or society. Extrinsic value will also depend on the context in which a product is generated because its value is dependent on its relation to other entities. In the context of new inventions usefulness might seem to require practical value. But this need not be the case. In the production of art, philosophy and many other fields practical applications of products may be hard or impossible to find nevertheless products may be useful in providing new thoughts and in changing minds. Boden (1995) suggests that in the case of creativity value is best understood as the generation of interest. Creative products are valuable if they are interesting. I will embrace interest as a way of ascribing value to creative products.

Extrinsic value or interest, like novelty, needs to be contextualised. To describe a product as interesting (or valuable) will require one of two strategies. Either interest will be a normative assessment dependent on the community within which it is created, or it will have to be an internal judgment by the creator. If the criteria are external to the creative individual, i.e. they depend on a judgement made from a social or historical point of view then they will not be limited to any particular way that creative products come about but depend solely on the particular normative criterion of interest. This is not surprising as interest or value is a very broad characteristic to ascribe to a product. It is novelty that is supposed to pick out the limits of which products are possible creative candidates. Any further judgment of interest or value will be dependent on an earlier judgment of novelty to pick out creative cases.

It will be unlikely that we can assess the agentive way that a grouping of creative products, based on the criteria of novelty and value, arise as a collective. For the range of products, grouped by an external judgement as novel and valuable, seem to come about from a variety of different creative phenomena. This is problematic for accounts of creative agency for it may be that understanding how our reasons are connected with the production of creative products will require explanations that are particular to the specific types of creative phenomena that produce them.

1.3 A Necessary Agency Clause for Creativity

One tactic for ensuring that creativity is linked to agency is to include a further necessary condition that for a product to be creative it must, along with being novel and valuable, be produced in an agentic way. One such account is presented by Gaut (2012). Gaut is interested in providing an account of creativity as a rational process. He suggests that the activity of producing creative products can be broadly considered as equivalent to doing any other activity of production. Just as when you prepare a meal you are guided by your reasons for making that meal, in the same way when you create art you are guided by your reasons for creating art. So for people to be rational in preparing a meal or producing creative products they must be appropriately sensitive to their reasons for carrying out that activity²⁰. However according to Gaut there is a disparity between carrying out general tasks and making works of art, that artists are often irrational, and that the very irrationality may be important in the production of creative products.

Gaut (2012) highlights the fact that psychological surveying of creative individuals has shown an increased likelihood of those individuals suffering from a mental illness or condition (see Andreasen, 1987 and Kinney et al., 2001). Gaut is motivated by the concern that the psychological evidence points to a link between creativity and irrationality but rather than accepting that creativity is fundamentally irrational or denying that a link exists between irrationality and creativity he tries to provide an account of rational creativity that can accommodate irrational episodes.

In Gaut's account, for a product to be produced in a rational way and therefore meet one of the criteria for being creative, it need not be produced by a process that is rational throughout. The rational creative process may include irrational events, creators may be beset with visions, caught up in the moment or be inspired by irrational desires or emotions. The only necessary rational aspect of the process is that the production of the product be sensitive to reasons. Gaut (2012) utilises the term product-value rationality. He suggests that what is important for ascribing agency to creative products is that at least part of the process of generating a creative product is guided by an individual's reasons for considering what they are producing as being valuable in a sense relative to the field in which the product is produced. We can suggest that what product-value rationality amounts to is individuals being sensitive to their field relative reasons for producing the product that they do, for behaving or acting in one way or another, within the particular context of producing that product.

In many circumstances we might consider that product-value rationality is enough to establish agency over the production of creative products. Even in cases where there is irrationality involved in the process, it might be the case that the process can be considered rational overall. In these circumstances Gaut is moved by the consideration that all creative processes require rational events but not all require irrational events. Therefore, for Gaut, the overall character of these processes is rational.

²⁰ This account of rationality addresses the question of creative agency for being reason sensitive can be seen as a condition for an individual to be morally responsible for a particular behaviour (see Fischer and Ravizza, 1998).

Will this characterisation of the rational nature of the creative process suffice as a way of providing an agentic condition for creativity? Well, there remains at least one common type of case that is problematic for Gaut's claim that all creative products are brought about in a reasons sensitive way: cases of insight. Gaut (2012) refers to cases of insight, where creative products seem to pop into the mind in an 'Aha' moment, as instances of spontaneity. For Gaut (2012) spontaneous events are of the same character as instances of pain. They can be attributed reasons in the sense that they have causal explanations but they are not normative facts and as such they should not be considered rational or irrational but *non-rational* in character, that is they are not sensitive to reasons. Gaut claims that such non-rational events will not affect the overall claim of whether a process is rational or not. Since they are non-rational rather than rational or irrational they do not affect the rational character of the process. Therefore a process can be rational, and thereby meet the agency condition of being creative, even if it contains non-rational events as long as it also contains episodes of product value rationality. But consider an example where the idea for a new painting springs into the mind of the artist in a moment of insight. How can this event be seen as rational, as guided, or sensitive to reasons? Gaut doesn't suggest that it is; he relies on other events to make this non-rational event part of an overall rational creative process. But this may be problematic. If we limit our consideration to the moment of insight then we can suggest that the overall nature of the process is not obviously rational. If these non-rational events as Gaut describes them are not guided by reasons then they will not be agentic, and if they can produce products that are novel and valuable, then we have reason to doubt that all creative products are brought about in an agentic way. Therefore Gaut's claim that all phenomena that are creative are also responsive to reasons seems premature. In the moment of insight we have a prototypical example of a creative phenomenon that can produce novel and valuable products but yet seems intuitively involuntary and non agentic (as Gaut suggest in his characterisation of insight as non-rational). Being sensitive or responsive to reasons then, even if it is a primary characteristic of agency, will not be applicable as a necessary agency clause for creativity unless it can accommodate IBPS. But, as yet, it is not clear here how it can be applied to that process. Understanding when and where IBPS is reason sensitive is one of the main goals of this thesis.

1.3.1 Understanding Creativity as Broadly Intentional

Another philosopher that suggests agency as a necessary condition of creativity is Søren Klausen. Klausen (2010) considers that such a condition should not be too constrictive. It should not require an agent to understand exactly what is engaged in. Consider Fleming's discovery of penicillin. Fleming was not aiming to discovering penicillin when it grew in his agar plate but was trying to discover something new. For Klausen this event may still be creative even although Fleming was not directly searching for antibiotics. Klausen (2010) is quite vague about what he considers to be the necessary agentic condition. His strongest hint is that an individual can only produce creative products if they intend to bring about something new. So if an individual wanted only to solve a problem through conventional means and

somehow found a solution that was unconventional, his discovery should not be considered creative if he had no awareness or intention to bring that about. Here we may think that Klausen's condition runs into the same difficulty as Gaut's (2012), for one common instance of how creative products come about is through IBPS. Here we can suggest that individuals are aiming to solve problems through normal means yet novel and valuable answers arrive that we don't intend and are unaware of producing. If Klausen's definition rules out cases of insight that produce novel and useful products because individuals did not intend to bring about new products then it will be untenable as a condition for a product to be creative.

Further the intentional or agentic nature of the creative process is here assumed to be unproblematic but this is, as we shall see in chapter 3, not the case for IBPS. Further this broad sense of intentionality may not lead us to a very informative account of creative agency. It may tell us little about how we are agentic in terms of creativity and creative products, only that it is something we aim at doing, in some vague sense of bringing about something new, and this as I have already suggested may just be presupposing that we actually are aiming to do something new (all the time) when we produce creative products, a claim that is far from being established. For we might think in the case of IBPS we are not aiming to produce novel products yet such products arise. Klausen's criteria then may fail for a large class of events we generally think of as creative. Accounts of how creative phenomena are agentic will have to do better.

1.4 Novelty and Psychological Creativity

So far I have suggested that because accounts of novelty pick out a broad range of creative phenomena, assessing the association between agency and this range of creative phenomena will be problematic. Instead we should address the question of creativity agency by assessing a particular creative phenomenon. I have suggested looking at IBPS. This has significant appeal because it addresses directly some of the difficulties associated with understanding creativity as agentic.

Some might think that this move is too hasty; after all many models of creativity have been suggested. I will not assess them all here. I will however look at transformational accounts of creativity for they have been highly influential. It will be worth our time exploring if this type of account can provide a definition of creativity that will be useful in addressing the question of creative agency. This will also provide some background to the debate concerning novelty and psychological creativity.

Boden (1990) provides the paradigm of a transformational account. Her account is an attempt to understand P-creativity. She suggests that the type of mental mechanisms involved in bringing about new ideas should be central to our consideration of whether the products of those mechanisms are sufficiently novel to be considered creative. Her approach is to provide a range of mechanistic explanations that suggest different ways novelty arises. Boden's account has been widely adopted. Stokes (2011) for one utilises a Boden-like definition of novelty in his account of minimal creativity.

In defining novelty in a mechanistic way Boden hopes to remove the need for a contextual judgement about whether or not a product is novel and in this way naturalise the ascription of novelty. If this move is successful then we may have a criterion of judging novelty that can be utilised to provide a unified concept of what is or is not creative. Such a definition might provide an account of creativity and therefore a stable base from which to assess the question of creative agency. If this is right we might think that investigating the agency associated with this definition of creativity would be more worthwhile than assessing the agency involved with IBPS. I will not suggest a knockdown argument against Boden's position but I will argue that focusing on IBPS rather than Boden's transformational account of creativity will yield better rewards for addressing the question of creative agency. In suggesting this approach I am not committed to the truth or falsity of Boden's position. But in examining Boden's account I am committed to identifying some difficulties with the account that will make providing an account of creative agency based on her definition problematic enough to justify an alternative approach.

1.4.1 Boden's Definition of Creativity

In her book 'The Creative Mind' Boden (1990) presents a distinctive definition of creativity. To achieve this she contrasts processes that produce p-novel product with process that produce p-creative products. P-novel products are new in that they have never been realized before by that psychological system. To generate p-novel products you need only utilise a mind-set that you already have in place. Take for example a machine that makes sweets in five different shapes and five different colours. You just push the buttons for the colour and shape you want and out comes the sweet. The sweet machine can make 25 different kinds of sweets. The first time you make a sweet that resembles a red car or any combination of shape and colour, the product is novel, for it has never been produced before by that machine. But since you always had the capacity to produce any of 25 sweet combinations it should not be considered creative, for creativity requires something more than mere novelty. In relation to p-creativity the product of a psychological system will be p-novel if it has never been produced before by that system. For Boden (1990), for a thought to be p-creative it must be novel, i.e. never have happened before, but it also requires a further condition that it 'could not have happened before'.

To understand what Boden means by 'could not have happened before' we need to understand her concept of *conceptual space*. Conceptual space for Boden is a mental property of a specific type of thinking. It is a representation of the rules or principles through which a unified process of thought specific to an area of consideration is generated (Boden, 1994, p79). Here the rules may not be specifically consciously accessible to the user but are bound up with how they conceive the problem. The conceptual space is however a representation of the rules. For example, our understanding of mathematics allows us to manipulate numbers in different ways. How many ways depends on the depth of our understanding of mathematical rules. These are the same rules we utilise every day when engaging in a mathematical task. Our understanding or representation of what tasks can be carried out within these rules is our

mathematical conceptual space. According to Boden (1990) for a thought to be creative it must, in being formed, transform the rule base under which we are currently operating. Cognitive processing that only utilises the rule base or conceptual space we already have in place is not creative because it does not transform that rule set (Boden, 1992). Such transformations can occur in a number of different ways, by the dropping of a rational constraint or by reversing a rule. By dropping a constraint the conceptual space supporting the enquiry is altered and new possibilities can be realised. Similarly, conceptual space can be transformed through the negation of a constraint (by considering the negative or reversal of a rule). In dropping one rule or another or by changing a rule in a pre-set manner, such as negation or reversal, it is possible to make connections and produce products outside of what our standard conceptual space could have achieved.

P-creativity then, for Boden (1990), is a matter of conceptual space being transformed. P-creative thoughts change the rule base through which you think and solve problems. They could not have occurred before by utilising the rule set that currently comprises your conceptual space but require a transformation in that thought process to bring the creative product about. P-creative processes are those processes that transform conceptual space. P-creative products are produced by such processes. In defining p-creative processes in this way Boden can make a distinction between p-novel products and p-creative products. A p-novel product is one that can be produced by our current rule set. A genuinely original or radically p-creative product is one that cannot be produced in this way; for such a thought to arise we must transform our conceptual space (Boden, 1990).

1.4.2 Exploring and Transforming Conceptual Space

Within her framework of conceptual space Boden (1990) makes a distinction between exploring and transforming such a space. Conceptual space can be explored in many different ways. In general, we are exploring conceptual space when we engage in exploring what the conceptual rules currently at our disposal are capable of doing. Novel possibilities can be uncovered by such rule based exploration. On other occasions exploring the conceptual rules can demonstrate to us the limits of our current conceptual thinking. In some cases such exploratory thought may identify points where conceptual transformations can be made in the future, whilst at other times it may be a wasted effort of churning over unhelpful possibilities.

Boden provides an example of the transformation of conceptual space in classical music. 19th century classical music utilised a strict structural element known as the home key. A home key is the key in which a piece starts and ends. The key is reinforced within the body of the work by accompanying chords or melodic scales. This formal structure provided a conceptual space for composers to explore. The increasing exploration of that space by Bach and other composers reduced the need of home key reminders and forced the structure to breaking point. Finally, Arnold Schoenberg dropped the home key altogether. In so doing, conceptual space was not just tweaked as it had been throughout the process but a

transformation of the conceptual space of composing classical music took place (Boden, 1994)²¹. The tweaking of the conceptual space through the increase in the range of home keys by Bach and the subsequent reduction in the need of home key reminders can be seen as examples of exploring the conceptual space of classical music at that time. Distinctly, Schoenberg's dropping of the home key altogether can be seen as a transformation of conceptual space through the dropping of a constraint (Boden, 1994).

1.4.3 Transformational Creativity, the Tripartite Model: Necessity and Sufficiency

Boden (1990) ties P-creativity to these strong cases of transformation, where conceptual space is transformed in producing a new product. Here it depends on our intuitions whether we should think the title P-creativity should be reserved for such strong transformation cases or whether it can be applied to novel cases of exploring conceptual space. Consider Boden's example of Bach. She suggests his exploration of the conceptual space involving home keys should be seen as a novel exploration of a conceptual space but not as an example of creativity. But to many this is counterintuitive. Bach's exploration may fit the standard requirements for novelty, value and agency without being transformational. In cases such as these are we not right to suggest that Bach was being P-creative. In light of such criticism Boden (1995) clarifies her position. She suggests that in presenting her work (1990) she was unclear and that her emphasis on *transformational creativity* was too strong. What she had in mind and what she presents in later accounts (1995, 2009) is a three tier conception of creativity. In this account transformational creativity forms the first and most creative tier. As in her 1990 account transformational creativity requires the generation of creative products through the altering of conceptual space. The second tier is termed *exploratory creativity*. Exploratory creativity does not require the transformation of conceptual space but occurs when novel combinations are arrived at by exploring within the existing boundaries of an individual's conceptual space. Boden thinks that important discoveries within this space are entitled to be called creative. Finally the bottom tier is called *combinational creativity*. Combinational creativity occurs when novel products arise through the combining of familiar ideas. Here, then, Boden presents a tripartite definition of creativity based on three separate creative mechanisms. Along with satisfying the conditions for the mechanism in question all products require a further condition to be considered novel, that is that they have to have never been produced before by that individual.

The tripartite model is interesting because it tries to provide us with three distinct mechanisms of how creative products come about. Boden (2009) does not presume to have captured the only three ways it is possible to produce creative products. She suggests the three mechanisms should be seen as providing

²¹ Within rule changing, this is an example of rule dropping.

three separate conditions for the production of creative products. If products are suitably valuable, are produced in one of these three ways and have never been produced before by that system, then this will suffice for them to be considered creative.

This reply to the criticism that transformational creativity was too strong raises some difficulties for Boden's definition of creativity. One of the benefits of transformational creativity was that it provided a straightforward way for considering a product to be sufficiently novel to be creative. They needed to have been produced by a system that could not have produced them before. The tripartite model removes this strong requirement. Instead we require a new way of defining which products will be sufficiently novel to be creative. Products will still have to be, as is the case in transformational creativity, never produced by that system before. But this will not be sufficient for a product to be novel enough to be considered creative for there are combinational novel products, like the car sweets, that we would not judge as creative even if they have never been produced by that system before.

Boden herself is also concerned with the issue of whether the mechanistic account fully captures the notion of novelty that is required for a product to be creative. One response to this problem presented by Boden (1995, 2009) is the idea that creative products should be in some sense surprising. Novel products would be seen as creative if they are generated by any of the three mechanisms and are valuable. Whereas novel products that are not surprising, like the red car sweet from our earlier example, should not be considered creative even if they were generated by one of the three mechanisms. Such a stipulation may have an intuitive appeal, in that we are often surprised by genuinely creative ideas but not by long novel number strings. But will the concept of surprise be specific enough to give us a good definition of what sorts of novelty will ground an ascription of creativity?

Surprise is not limited to cases of creativity. We can be surprised by our angry response to a television commercial even if such a response is in no way novel. Surprise then will act as a further condition for creativity. For products to be novel enough to qualify as creative it will be sufficient for them to be produced by one of the three mechanisms, be valuable and be surprising²².

But it does not seem impossible that there will be cases of creativity that are perhaps unsurprising to the individual but nevertheless be genuine instances of creativity. This will depend on how we understand surprise. Surprise or the property of being surprising can be understood as an external judgment or as an internal judgment by the individual producer. If surprise is considered an external judgment then we might be concerned that it does not capture the P-creative context. For example

²² Boden (2003) suggests a strong notion of surprise in an attempt to capture the strong sense of creativity she had originally outlined in 1999, that creative products should not have been able to be produced before. The notion of surprise concerns being surprised not by what happens but by the possibility that the particular event could happen. A condition of creativity on these lines may be too strong; why should creative products have to be surprising in this fashion.

consider an expert assessing a composition by a young musician. The work although novel to the musician will not be surprising to the expert if the variations involved are ones commonly utilised by advanced musicians. This may be an indication of the talent of the student but it will not be particularly surprising to the expert. An external claim of surprise, then, is unlikely to capture the P-creative context and therefore as a necessary condition for p-creativity it seems misplaced.

What about an internal judgement? An internal judgement of surprise will be associated with the P-creative context. Here surprise will have to be understood as a judgment or reaction of an individual to the phenomenology of a specific mental event. This experience may be too individualised. What might be surprising for a non expert may not be surprising for an expert even if it is novel. This may make an ascription of creativity based on an internal judgement of surprise problematic for we might consider the work of the expert to be creative even if he doesn't find it surprising. A requirement that an individual finds their creative products surprising may not be strongly enough associated with novelty to make a particularly good way of separating products that are novel enough to be considered creative from those that are not.

Overall the tripartite model does not pick out all the different ways we can be creative. It is a mechanistic account that suggests three ways that are sufficient for producing creative products. Boden is not supposing that she has captured the only ways we can be creative but identifies some important mechanistic ways that creative products are produced. This difficulty in providing necessary conditions under which products should be seen as an instance of P-creativity is itself not surprising because the tripartite model is based on novelty and value; its aim is not specifically to produce necessary conditions for creativity but to improve our understanding of creativity in a subject neutral way by which particular products, for example the products of computers, can be judged.

1.5 Naturalising Creativity

One response to the claim that the concept of creativity is associated with too many different phenomena for the relation between creativity and agency to be properly assessed is to suggest that some of these phenomena are superficial. Even if we intuitively think there are different creative phenomena there may be a specific type of cognition that underlies and is fundamental to all creativity phenomena. I will, in this section, outline the main themes of research into the psychology of creativity. I will suggest that as the research currently stands it most likely suggests that the concept of creativity is underpinned by a wide variety of cognitive processes.

Boden's account of transformational creativity can be seen as one attempt to provide a unified mechanism for generating creative products. But how natural and therefore how useful is Boden's or similar transformational accounts of creativity?²³ One criticism of Boden's account is that it is too vague

²³ For a similar transformational account see Wiggins, 2001.

to be of use in mapping actual brain processes. The following discussion is not intended to be a knock down argument against the usefulness of utilising a transformational definition of creativity for understanding creative cognition; rather it is my hope that highlighting the difficulties related to such a definition will provide some justification for adopting a more piecemeal approach to understanding creative agency.

1.5.1 Can Transformation Separate Creative Cognition from Basic Cognition?

One major difficulty with definitions of creativity based on transformation is separating creative cognition from basic cognition. Boden considers cognitive events only to be transformations if they could not have happened in the system before that point in time. The distinction between basic and creative cognition puts great pressure on the concept of conceptual space. There are a number of objections as to both the coherence and the usefulness of this notion (for a range of objections see Ritchie (2007); Perkins (1995); Ram et al. (1995); Schank and Foster 1995; Turner, 1995).

Boden (1992) suggests a traditional four stage approach for modelling creative cognition²⁴ within which conceptual space is searched, broken and redefined (Boden, 1990). As Boden has stated, not all creative products will require such a transformatory process. But Boden considers the most novel cases of creativity will be generated in this way. If transformation cannot be clearly distinguished from basic cognition then we have good reason to think that Boden's account cannot be naturalised.

One way of supporting a claim of transformational creativity would be to empirically support the concept of conceptual space. Unfortunately conceptual space seems particularly hard to investigate empirically. The best evidence would be direct biological evidence, but here we encounter the difficulty of distinguishing basic cognition from creative cognition, in as much as monitoring changes in brain behaviour will not tell us if those changes are transformational or not. Nor will they tell us whether they are combinational or exploratory. If we wish to stick with identifying p-creative episodes than we might suggest that the reports of the individuals undergoing experiments should be investigated in order to try to identify underlying mechanisms. Such reporting is beset with the usual difficulties of accuracy and reliability. But more importantly trying to pick questions that will be able to be accurately answered is quite difficult. Did the answer transform your conceptual space or merely change it? Answering such a question will be very subjective. Individuals will feel differently about similar experiences. Consider the following insight problem. You are presented by a piece of paper 3 inches by 5 inches (see appendix 1 0.2). How can you cut a hole in the paper that you can put your head through? See appendix 0.3 for the solution. In finding the solution how much did you change your conceptual space concerning the problem? Did you just tweak it or transform it? The same experience by two individuals may be described in different ways either as transformational or as a substantial change or just a change.

²⁴ Footnote: Four stage approaches can be traced back to Poincare (1908) through Wallace (1926) and are present in many accounts of creativity. See also section 2.4.

Furthermore the condition of surprise has similar difficulties. Do experiences of surprise correlate strongly with transformation? It may be that some people tend to experience surprise more readily than others and this may be entirely unrelated to the underlying mechanisms involved in producing the creative product or not correlated strongly enough to pick out the differences between transformational cases and non transformational cases. Such difficulties in measuring changes in conceptual space make it difficult to investigate scientifically. It may be better to avoid an investigation into the agency of transformational creativity because of the uncertainty surrounding its cognitive underpinnings²⁵.

1.5.2 The Psychology of Creativity: Six Ways to Study

We can now turn to consider the general field of the psychology of creativity. Since Guilford's (1959) address to the American Psychology Institute in 1959 there has been an increased interest in the psychology of creativity. Mayer (1999) characterises six different methodological approaches: *psychometric, experimental, biological, computational, biographical, and contextual*. I will briefly outline each methodology and suggest why they do not provide a straightforward definition of creativity that will be useful in addressing the question of creative agency.

1.5.2.1 Psychometric Approaches

Psychometric approaches view creativity as a mental trait that can be measured through appropriate testing. These tests start with Guilford (1959) and his investigating creativity via *divergent thinking*. Divergent thinking, the ability to come up with different ways of approaching a task, has been a trait extensively tested for, and the standard Torrance test was introduced by Torrance in the 1960s and 70s (Torrance, 1988). The Torrance test is a battery of 10 different tasks, seven of which are verbal and three figural (Crockenberg, 1972). The verbal tasks include the ask-guess task in which subjects are presented with a picture composed of an elflike figure staring at his reflection in a pond and asked to compose a list of questions about the picture (see appendix 0.3). A further question is then asked on the possible causes leading to the scene (Crockenberg, 1972). The second verbal task is the product improvement task, where

²⁵ One possible approach to save the notion of transformation is to ascribe novelty based on an external judgement of whether the production of a product transforms the social and intellectual context within which it is produced. Such an approach has been suggested by Ritchie (2006). This approach is controversial as it moves the context away from the psychology of the individual and so abandons one of the main aims of Boden's account. However useful such an approach will be for identify of products as novel it will be dependent on current understanding of the intellectual context within which the product is produced and not based on the way the products are produced. As I have already suggested a judgment of novelty and value cannot be utilised to separate the different creative phenomena. Now, this might not be important if all we want to do is provide a definition of creativity based on novelty and value. But on this definition of creativity any judgement about agency will run into the now familiar difficulty that it will have to be applied to all creative phenomena that can produce novel and valuable products. In particular the intuitively involuntary process of IBPS may be problematic.

subjects are asked the different ways they can think of improving a toy elephant to make the toy more fun for children (Crockerberg, 1972). The third task is the 'unusual use' task, where subjects are asked to list unusual uses for a cardboard box (Crockerberg, 1972). Four other tests fill out the verbal battery. The figural tasks include the picture construction activity where subjects are asked to draw something clever and unusual using an egg shaped piece of paper, the circles and shape task, the creative design task and the incomplete figure task. All the ten tasks are marked for flexibility, fluency, originality and elaboration (Crockerberg, 1972). An overall composite score called the creativity index can be generated by pooling the test scores, but Torrance discourages such a use preferring to understand the tests as highlighting diverging strengths and weaknesses within individuals (Kim, 2006).

Torrance (1974) suggests five areas where the tests may usefully be applied. First the test may help to understand the function and development of the human mind. Second, they can be used to uncover appropriate teaching approaches aimed at specific individual subject's needs. Third, they may be important in providing insights into appropriate remedial or psychotherapeutic programs. Fourth, they might be used to test the effects of educational techniques, teaching programs and course material. Finally they may provide a way of uncovering latent potential in tested individuals (Kim, 2006).

The test set then provides a standardised way of testing for individuals' ability at a range of tasks pre-designed to be indicators of creativity. Underlying such testing is the assumption that the specific tasks themselves are actually picking out to some degree the creative ability. Furthermore, scoring the tests requires that experts rate the tests for flexibility, fluency, originality and elaboration; therefore, within the testing, a set of evaluative judgments are required. Although the tests may be useful in individual and educational development they provide only a loose measure of actual creative ability. The psychometric approach is dependent on preconceived notions of what creativity is and on expert judgments of what is or is not original. Such a methodology therefore can only separate actual episodes of creativity from non creative episodes on this preconceived basis. Alone such judgments will not help us to define a specific creative phenomenon that we can scrutinize.

1.5.2.2 Experimental and Biological Methods

Experimental methods on the other hand focus on uncovering the cognitive process that underlies creative thinking. Supporters of this approach have tried to establish difference between creative cognition and non-creative cognition. Some examples of areas of experimentation are Sternberg and Davidson's (1992) investigation into the nature of insight; Runco and Sakamoto's (1996) work on the effects of brainstorming on creative production and Nickerson's (1999) investigation into the teaching of creative thinking skills. Kinney et al. (2001) tested for creativity ability in schizophrenia patients and family members. Biological methods also seek to uncover the underlying physical mechanisms that support creative thinking. They focus on measuring the bodily and brain changes associated with creative thinking. For example Martindale et al. (1996) demonstrated that creative inspiration may occur when attention is

defocused. In the last ten years a growing body of work has focused on identifying creative cognition through neuro-imaging studies (for overview see Dietrich and Kanso (2010) and subsection 2.3.8).

Both experimental and biological methods largely rely on the same methodology. Individuals are subjected to creativity testing and are analysed for social, psychological, biological or environmental differences. In this way it is hoped that underlying differences in these domains can be identified with individuals who display similar creative traits. So, for example, IQ test results have been compared to results from creative tests. Creativity shows an inverted U shape distribution with the highest levels of creativity being associated with individuals with high but not exceptional high IQ ratings (Kershner and Ledger, 1985). The hope then is that such testing can help identify psychological, social, environmental, and biological differences that can shed light on creativity. One difficulty with such a method is that it relies on the strength of the creativity tests. Just as in the psychometric testing above we may be unsure as to whether such tests can be identified with creativity. Scoring highly on such tests should only be seen as an indication of creative ability; it should not be seen to be identical with it. So at best such comparison experiments can only provide indicators of underlying correlations with social and environmental factors. The same is true for both psychological and biological extrapolations.

However this type of work does provide a window on creativity. To date what has been indicated by the research is a large array of brain activity and a large number of social, environmental and psychological factors that can be linked with creativity but no unified understanding (Csikszentmihalyi, 1999 and Dietrich and Kanso, 2010). This highlights the difficulty for understanding the agency involved with creativity, for even if we gather a lot of evidence that correlates with creativity it does not necessarily give us an indication of a single unified phenomenon which can be assessed. The most likely explanation for the current data is that the varied testing base picks out an array of biological underpinnings because the target of the tests, whatever activity is picked out by the standardised testing, does not map to specific brain operations but to a wide array of cognitive abilities (Dietrich and Kanso, 2010). So such testing supports the hypothesis that creativity cannot be understood as a single phenomenon. It may be, then, that to address the question of creative agency we will need to focus on creative phenomena that have good cognitive underpinnings.

1.5.2.3 Computation Methods

Computational methods take their starting premise from the idea that human creativity can be modelled on the information processing of computers; that by coming to understand how computers might be creative we could have important insights into how we are creative. Margaret Boden (1990) is an example of a researcher who believes computer modelling can help us to understand human creativity. Computational methods may provide a way of modelling the creative process as a whole. Above we looked at Boden's attempt to do just that (section 1.4 and subsection 1.5.1). Boden's particular approach did not provide an accurate way of separating creative cognition from non-creative cognition. On the whole, computational

models tell us something about how theoretically creativity might work, that is the best way a particular problem can be solved given our current understanding, and this may go some way to informing how we think our brains do work. But if Boden's attempt is suggestive it will most probably point to a large number of different creative phenomena for it is unlikely that a single cognitive model will be able to accommodate all the possible ways a novel and useful product can be produced. As such the question of whether we are agnostic in regard to creativity will depend on addressing creative phenomena individually.

1.5.2.4 Biographical Methods

Contrary to the above methods biographical methods use case studies of prototypical creative people as a starting point for investigation. Here the individual's own account of their creative experiences and historical evidence surrounding the events are taken as evidence for particular accounts of creativity. This approach was characterised by leading figures in the early investigations into creativity; researchers such as Cox (1926) and Galton in his 1869 'Hereditary of Genius'. A more modern example of this approach can be found in Wallace and Gruber (1992) including their own examination of Charles Darwin. Other studies include Albert Einstein by Miller (1992) and Faraday by Tweney (1992). The methods involved here depend on the reliability of the reporting both from the creators themselves and other historical documents. One difficulty arises from the selection of the cases to be studied. The criteria of selection have often focused on examples of H-creativity, so such investigations have tended to lead to the reinforcing of the idea that creativity as a process is available to only a relative minority of individuals (Mayer, 1999). A second difficulty arises from the fact that the focus of biographical investigations is often on the experience the individual under investigation has of the creative process. Artists have been prone to overplay certain aspects of the creative experience to reflect preconceived conceptions of creativity and the 'proper' nature of the artistic process (Mayer, 1999). Individual reports on the creative experience may therefore be misleading. The difficulties involved in relying on reports and prioritising pre-selection criteria make this approach unlikely to provide a particularly clear account of the phenomena involved. Phenomena that are suggested by this approach will have to be investigated by other means if a scientifically credible account is to be forthcoming.

1.5.2.5 Contextual Methods

Contextual methods focus on understanding creativity in its social and biological context. An example of such research was Lubart's (2000-2001) examination of how the cultural environment affects creativity by comparing creativity in different countries. Again such experiments will rely on one of two methods for accessing creativity: evaluative judgments of novelty and value or psychometric testing. As such although they may provide interesting findings on the interrelation of creativity and the environment they are

unlikely to provide a unified definition of creativity cognition that will be useful in addressing the question of creative agency.

1.5.2.6 Conclusions on the Psychological Approaches to Creativity

The differing nature of the research on creativity complicates the field, for often in each methodology different notions of creativity are being used and it is therefore difficult to understand how results should be understood in relation to each other. But in trying to understand the overall category of creative cognition these difficulties might be overcome. These questions are not addressed in this thesis. What I am interested in is whether creativity is agentive. We have two possibilities for investigating this topic. Possibility one: we can assess the concept of creative as a whole; to do this we need a unified understanding of how creative products come about. As we have discussed above, no such account looks likely. Possibility two: if creativity is dependent on an array of different phenomena we can opt to assess each separately. My approach will be directed by the assumption that this is the case. The range of creative phenomena may be large and ill-defined. I will focus on IBPS because it presents a particular challenge for understanding creative agency but before I do that I will provide a justification for focusing on IBPS rather than some other creative phenomenon.

1.6 Creative Problem Solving

One further attempt to frame creativity as a single phenomenon is to consider it in terms of problem solving. A wide range of psychological literature (for an overview see Weisberg, 2003) suggests that a credible body of what we normally understand as creativity can be successfully understood as a type of problem solving (for a philosopher who also agrees with this view see Briskman, 1980). Weisberg suggests that any solution to a problem that is novel and fulfils the requirements of the task can be considered creative. Some (Lubart, 2000-2001) have criticised this view by claiming that problem solving does not cover all types of creative phenomena. This objection may have some intuitive support; we might consider that there can be processes that produce both novel and valuable products that cannot be modelled on problem solving. One common example is sketching. In sketching an individual may come up with creative products without having engaged in any of the elements of problem solving (Lubart 2000-2001). For our purposes this may not be a difficulty. If creative problem solving picks out a particular kind of creative phenomenon then focusing on creative problem solving may be a useful way of addressing the question of creative agency. One difficulty of this approach is finding a good definition of problem solving and finding a model that can be agreed on.

One particular division between problem solving models concerns models that do and do not involve insight (Sandkuhler et al., 2008). IBPS is thought by many to involve distinct cognitive processing to NIBPS (Sandkuhler et al. (2008) and see section 0.4 and chapter 2 for a full discussion). The category of creative problem solving may be divided by underlying cognitive mechanisms. If this is the case then

any specific claims about the agency of creative problem solving as a whole might be undermined by these different cognitive underpinnings. IBPS may present particular difficulties for the question of creative agency since insight seems to be intuitively involuntary and non agentive. Thus utilising creative problem solving as a category for addressing the question of creative agency may be problematic.

Even when linked with problem solving creativity remains too vague as a topic of enquiry. It allows for a wide range of possible creative phenomena. In light of this the best approach will be to analyze particular creative phenomenon that are well grounded in the scientific literature. I will focus on the phenomenon of IBPS and in the next chapter I will present the scientific case for understanding the nature of IBPS as a discrete type of creative phenomenon with substantial scientific support. Before then I will discuss the relation between IBPS and creativity and suggest why understanding IBPS will be particularly important for the question of whether creativity is agentive.

1.6.1 Where Does This Leave the Concept of Creativity?

If we want to understand the association between IBPS products and creative products we will require a definition of creativity. In this context creativity maybe best understood as c-creativity (see section 1.2.1). For products to be creative they will have to be judged as novel and valuable within the context in which they are created. This leaves open the debate about how we should define these conditions and the nature of any agency condition. The best approach to understanding novelty might be to gauge it by an evaluative judgment. It may require that products have never been produced by that person before, but beyond that any definition of novelty will depend on a context based judgment. Value might best be seen as the generation of interest, an approach advocated by Boden (1990, 1995). If a product is interesting then it has some value associated with it. Interest will also be context specific, so products may be interesting and therefore valuable in some instances and not in others. Again it might be captured best by an evaluative judgment based on the context in which the product is produced. These definitions will only be used as guidelines for discussing the relation between IBPS products and creative products. It is hoped that it will be inclusive enough to assess how it is possible to understand IBPS products as creative products. We will examine this association fully in the conclusion after the agentive nature of IBPS has been established. However, we can at this time begin to discuss the association between IBPS and creativity, and in so doing we can provide a justification for examining IBPS.

1.6.2 IBPS and Creativity

The phenomenon of the moment of insight has been included in many historic, almost mythical moments, of human invention: Archimedes jumping into his bath tub; Newton being hit by the apple. The historical accuracy of these episodes aside, the inclusion of the insightful experience, that sudden coming to mind of new ideas in these episodes of revelation has perhaps forever linked creativity with this eureka experience. Nor does the connection of the two phenomena stop there, for insight seems to happen not just to historical geniuses at moments of great importance in human history, it seems to happen to all of us. Most

of us will admit to having experienced what has come to be known as the ‘Aha’ experience, where the answer to a previously toiled over question just seems to pop to mind. This experience is a central motivating factor for an investigation into creativity. It has presented a mystery to be understood. Investigators into creativity have often focused on the eureka experience as a way of understanding the concept of creativity as a whole. But, as we have been discussing, limiting creativity to a singular phenomenon seems fraught with difficulty²⁶. No attempt then will be made to understand creativity solely as IBPS. Rather IBPS will be investigated as a prototypical example of a creative phenomenon. We have defined creativity without the inclusion of a process. Being an IBPS product is not then necessary or sufficient for being a creative product. There is however a strong link between IBPS products and being novel and valuable. IBPS products often lead to a problem being solved; in that sense, most of these products are valuable. In terms of interest we might consider nearly all solutions valuable in that they often provide new ways of looking at a problem. IBPS products are also often p-novel (subsection 1.2). IBPS involves providing answers to problems that have not been solved and perhaps could not have been solved by straightforward means. They are often generated after an impasse in cognitive thinking where no solution is available (Beefink et al., 2008). Answers to problems provided by moments of insight are therefore often outside the range of the obvious solutions to problems. In the psychological context of the individual that produced them IBPS products will be predominantly judged as novel. IBPS, then, is a cognitive mechanism that has a propensity to produce products that will be judged to be novel and valuable; as such, understanding the agentic nature of IBPS is vital for understanding how creativity is associated with agency and any acceptable definition of creativity has to be able to include the phenomenon of IBPS.

1.6.3 The David Objection: What About an Agency Clause?

In providing the above definition creativity can occur by any means. The statue of David, (section 0.21) if judged as novel and valuable, will be an example of creativity even if it was formed by rain weathering limestone. This to most of us is not intuitive. We might think that in pursuing an agentic account of IBPS based on reason responsiveness we can include a necessary agency condition for creativity based on those accounts. Agency would still act as necessary condition for creativity, but ascribing agency in this context would require investigating the different creative phenomenon that lead to creative products.

This is perhaps not the best approach. Creativity is a very broad concept. The range of behaviour that different people consider creative suggests that it might be useful to be flexible in our assessments of the agentic nature of creativity. For instance we might consider that the activity of animals has a claim of being creative. Consider the way particular groups of bottle nose dolphins hunt for fish on sand bars (Shane et al., 1986), or an Orca hunt for seals in sheltered inlets (Lopez and Lopez, 1985) or the way

²⁶ Stokes (2007) suggests that insight should not be considered as identical with creativity. He utilises the term incubation essentialism to refer to accounts of creativity that depend on incubation.

Bonobos gather ants (Sanz et al., 2009). All might be considered creative. So there may be a range of behaviour that produce novel and valuable products and that come close to being agentive and therefore are on the cusp of what we might consider genuine examples of creativity where intuitions are divided about whether a specific activity qualifies. A good example of this is the question of whether computers or robots can be creative. Some hold the opinion that computers or robots can be creative or at least hold the intuition that this is quite possible even if no truly creative robots or computers have been made by us (Boden, 1990; Bird and Stokes, 2006). Others hold the intuition that robots or computers cannot be creative (Gaut, 2010). So in defining creativity it might be best to adopt an inclusive policy that leaves room for other types of agentive clauses to be developed.

One criterion that clearly ascribes agency to episodes of behaviour is being reason responsive. We can suggest that being brought about in a reason responsive way is sufficient for a product that is novel and valuable to be creative. This does not rule out other ascription of creativity based on different standards of agency but it provides a standard that if it can be met will satisfy our intuition about the need for creative product to be produced by an agent. For products to be creative then it will be necessary that they are novel and valuable (see subsection 1.6.1) and it will be sufficient that they are produced by a reason responsive process. But to rely on reason responsiveness as a condition of creativity we need to be able to understand IBPS as agentive for as we have discussed IBPS is a prototypical creative phenomenon that often produces novel and valuable products. Providing an account of reason responsiveness for IBPS is the goal of this thesis. In providing this account we will be able to add to a definition of creativity the condition that for a product to be creative it is sufficient that it is reason responsive.

1.6.4 Possible Dangers of Focusing on IBPS

One objection to focusing on IBPS is that it is too narrow a topic; to understand creativity as a whole we need to look beyond IBPS. In focusing on IBPS I am focusing on a particular creative phenomenon. I am doing so because I want to look at the agency of creativity and the particular problems associated with understanding IBPS as agentive. In so doing we can address the difficult question of how creativity as a whole can be agentive when IBPS is a prototypical creative phenomenon yet is intuitively involuntary and non agentive. IBPS has a long history in the exploration of creativity. To some its importance has been overplayed (Weisberg and Alba, 1981). I have already addressed the way IBPS relates to our understanding of creativity as a whole (subsection, 1.6.2). Some specific points remain to be clarified. In IBPS the process is understood as an individual pursuit. That leaves to one side the importance of social and environmental interactions. The thought that creativity does not go on in isolation is quite right; the environment, cultural or otherwise is important in the production of creative products (Amabile et al., 1986). In focusing on IBPS I do not mean to suggest otherwise. There is a body of important work on creative groups and there may be some interesting ties with group agency (Maier, 1970; Paulus and Yanng, 2000; for a review see Paulus, 2000 or Hülshage et al., 2009). IBPS can be seen as one type of

phenomenon that has an important association with creativity; it may even play an important role in group creativity. I will not explore these associations but that is not to suggest that they are not useful ways of understanding creativity.

It can also be suggested that IBPS is disproportionately associated with particular fields. It has been seen to be particularly important in mathematics and science. It is seen as less applicable to the arts, where there are not definite problems to be solved. This may just be a matter of prejudice. IBPS may have an important role in nearly all creative fields. Problems arise in the works of artists just as in the works of scientists. Consider a novelist trying to resolve plot problems, or an artist stuck on how best to express the way a body moves utilising paint. Although all creativity may utilise many processes and techniques, IBPS is not exclusive to any particular domain. Particular examples of utilising drawing techniques as problem solving mechanisms or live action techniques in dance are perhaps best understood in a different way than problem solving or IBPS. In focusing on IBPS and defining it in a cognitive sense, these phenomena will require a different explanation and they deserve one. One of the advantages of focusing on IBPS as an individual creative phenomenon rather than creativity as a whole is to avoid bundling all creative phenomena into one artificial category. With the reasons for and the scope of my investigation into IBPS clear we can focus on the phenomenon of IBPS and investigate its current scientific standing. This will provide the proper basis for an accurate assessment of the agency associated with IBPS.

Chapter 2: The Science of Insight

2.0 Introduction

In our everyday lives we encounter a wide array of challenges and problems. Some problems can be solved in a straightforward manner. On the other hand some problems cannot be solved so directly; sometimes no solution can be found even after considerable thought. The answer to such problems often comes to us in moments of insight, out of the blue, after we have stopped deliberating about the problem at hand. I will call this type of problem solving insight based problem solving (IBPS). Providing a scientific understanding of this phenomenon is the central goal of this chapter. In what follows I will outline an intuitive model of IBPS that can be understood as a *four stage model* and provide scientific backing for this account. This account of IBPS will be utilised in the rest of this thesis to explain the activity involved with insight and to understand the question of creative agency as it is addressed to IBPS.

2.1 An Intuitive Account of IBPS

A common observation made by creative individuals is that creative solutions often seem to come out of the blue. They often occur when we are thinking of other things or nothing much at all. IBPS is associated with a particular reportable phenomenon. After the problem solver's deliberation has been unsuccessful in producing an answer to the problem at hand, a moment of revelation occurs in which the answer to the long pondered question pops to mind. This "popping" is a mental episode with the particular phenomenal character of a surprise revelation that answers, or seems to answer, the problem at hand. Within the psychological and philosophical literature it has been termed the 'Aha' or eureka moment. Here is Poincaré's recounting of that moment of revelation:

At the moment when I put my foot on the step the idea came to me, without anything in my former thoughts seeming to have paved the way for it [---] I did not verify the idea; I should not have had time, as, upon taking my seat in the omnibus, I went on with a conversation already commenced, but I felt a perfect certainty (Poincaré, 1908 pp 88).

The overall experience of IBPS is captured by Poincaré later in the same passage.

Returned to Caen, I meditated on this result and deduced the consequences [...] There was one however that still held out [...] But all my efforts only served at first the better to show me the difficulty, which indeed was something. All this work was perfectly conscious. Thereupon I left for Mont-Valérien [...] One day, [...] the solution of the difficulty which had stopped me suddenly appeared to me [...] I had all the elements and had only to arrange them and put them together. So I wrote out my final memoir at a single stroke and without difficulty (Poincaré, 1908 pp 89).

Wallace (1926) provides a four stage model of the phenomenon described in Poincaré's report. In stage 1, *preparation*, an individual pursues the problem in a traditional manner, evidence is assembled and assessed, and the individual deliberates on the problem at hand. In IBPS no answer is forthcoming and we enter an impasse where further efforts seem unhelpful. Here we enter stage 2, *incubation*. In incubation, unconscious²⁷ processes take over and provide a possible solution. Here it has often been thought that unconscious restructuring of the problem takes place that makes an insight breakthrough possible. In stage 3, *revelation*²⁸, the insight becomes available to the problem solver in an 'Aha' or eureka moment. Finally in stage 4, *verification*, the insight is assessed. Further work may be required before a final answer may be arrived at²⁹.

The four stage model provides the skeleton for an intuitive model of IBPS. In this chapter I will defend a particular conception of IBPS that is based on this model. Much of the work of the rest of the chapter will be to examine the scientific data on insight and provide a detailed account of IBPS that is in line with that evidence. Before providing this account it is important to show that IBPS is a genuine form of problem solving. I will therefore, first of all, review the evidence for and against IBPS as a distinct problem solving process.

2.2 Four Features of IBPS

The phenomenology of the 'Aha' experience has been a motivational force in categorising insight based problem solving as a distinct process. The process has been envisaged as a way of solving difficult problems that is accompanied by a sudden 'Aha' experience, and imparts important and memorable breakthroughs (Davidson, 2003). We can reinforce these observations by providing some common features that separate IBPS from non insight based problem solving (NIBPS).

Sandkuhler and Bhattacharya (2008) provide four salient features of IBPS that separate it from NIBPS. Firstly it is likely that individuals encounter a mental impasse; that is, they feel that their current efforts towards finding a solution are unproductive. This can be understood as the exhaustion of the possible solutions that are searchable under the current understanding of the problem. Secondly it is often thought that individuals must alter their mindset before they can proceed. *Restructuring* has been used as the term that represents this process. Through restructuring, the search space is altered; this allows for an insightful breakthrough and in turn problem resolution. Currently there are two opposing views of how restructuring takes place. These are the 'business as usual' view and the *special process view* (Bowden et al., 2005). In the *business as usual view*, there is no processing difference between IBPS and non insight based problem solving. All problems are solved using the same process. Reportable processes are active in

²⁷ Here unconscious is left as a non specific term. It indicates the fact that incubation is not accompanied by phenomenal experience and so is not reportable. No substantial theoretical claim about the nature of consciousness is being made here, or in this paper.

²⁸ The word illumination has also been used to denote this stage.

²⁹ The four stage model can be seen as operating in a number of repeating steps in order to provide a final solution. For a discussion see Lubart (2000-2001).

resolving problem tasks and they proceed via small incremental stages until a final solution is reached. In the opposing view, 'special process', non reportable processes are important in restructuring the problem so a solution may be found (Bowden, et al., 2005).

The third characteristic of IBPS as seen by Sandkuhler and Bhattacharya is depth of understanding. They suggest that IBPS has often been described as uncovering a deeper understanding of the problem than non IBPS. I will not emphasise this feature of IBPS. It seems intuitive that IBPS can lead to a deeper understanding but does not necessarily do so. Finally they suggest that insight is often described as arriving suddenly and without warning in an 'Aha' experience.

This leaves us with the following picture of IBPS. In NIBPS the goal and the possible means for achieving that goal are unambiguous to the problem solver (Naples et al., 2003). These problems can be solved through heuristic searches based on an individual's current understanding of the problem (Naples et al., 2003). In IBPS the situation is different. No solution is possible via a heuristic search based on how the problem solver currently understands the problem. Further, no representation of a problem solution is available to the problem solver. Utilising the current understanding of the problem leads to an impasse. For an individual to proceed, a *restructuring* of the problem must occur. On the special process view restructuring is thought to occur during incubation. It is widely thought to be a non reportable process through which the current way of understanding the problem is altered (Bowden et al., 2005). From this restructuring an insight is gained in an 'Aha' experience that leads to a new solution space being created within which possible problem solutions can be searched for. The view of IBPS outlined above fits the four stage model. The period of initial reflection and impasse *prima facia* matches up with preparation. Incubation cannot be phenomenally detected but can be supposed to occur because insight arrives suddenly out of the blue. This arriving in an 'Aha' moment is the insightful understanding and matches with revelation. The further work required to find an answer matches verification. It will be the work of the rest of this chapter to examine the scientific evidence for this intuitive account.

2.3 The Special Process View is Essential for IBPS

All the features noted by SandKuhler and Bhattacharya (2008) have been reported in connection with IBPS. However when we considering whether IBPS is a distinct cognitive process the nature of the restructuring process (incubation) is the most important factor. That is, in considering the claim that IBPS and NIBPS are distinct cognitive processes that differ not just in an individual's ability to consciously access but that also differ fundamentally in function, the two processes must resolve problems by utilising different processes. In the terminology we utilized above, if IBPS is to be a functionally different process from NIBPS then the special process view must be correct, the business as usual account states that all problems will be solved utilising the same means.

We can note that the special processes view might be preserved without the retention of the specific way it is conceived to function above. This will become important for some evidence against the

special process view will only count against specific features of that account. Further, in utilising the term special process as it has been adopted within the literature we are not committed to a process that is unique to IBPS. What the special process view depends on is a functionally distinct type of processing for IBPS and NIBPS. IBPS may rely on general purpose cognitive processes just as NIBPS but for the special process view to be correct they must differ in the processes that they utilise. I will proceed by examining the case for the special process view.

2.3.1 Evidence For and Against the Special Process View

The experimental evidence for and against the special process view comes from four main sources. The first is *cluster analysis* of insight problems, where convergence between ability to solve insight problems is measured against those individuals' ability at solving non-insight based problems. The second source concerns studies that look at whether verbal reporting, while engaged in problem solving tasks, has a disproportionate negative effect on IBPS over NIBPS. Thirdly, *feelings of warmth* experiments have been carried out that seek to measure whether there is a difference in individual reports (mid task), of how successful their problem solving is proceeding, between IBPS and NIBPS. Fourthly, fMRI, EEG and other neuro-imaging studies have been carried out to try to identify differing neural activity during insight and non-insight tasks. I will look first at the evidence for and against the special process view. I will then proceed by presenting the evidence for each of the stages of the four stage model. Although the picture is not entirely uniform and some evidence has been interpreted as supporting both the business as usual account and the special process account, a strong case can be made that IBPS utilises distinct brain processes over NIBPS and therefore that IBPS can be seen as a distinct type of problem solving. Further there is substantial support for the existence of each of the stages of the four stage model.

2.3.2 Early Evidences Supporting the Business as Usual View

Weisberg and Alba (1981) set out to test the role of *fixation* in two insight problems: the nine dot problem and the triangle problem. Fixation is the idea that problem solvers can become locked into a search space that cannot yield an answer. It was seen as a reason why certain types of problems produced insights, for after fixation a problem may require an insight or restructuring before it can be solved; see Davidson (2003) for an overview of early thoughts about insight. In the nine dot problem you are required to connect all nine dots with four lines without removing your pen from the page (turn over for solution). In the triangle problem you are required to form four equilateral triangles using only 6 match sticks (see appendix 1 section 0.4 for solution).

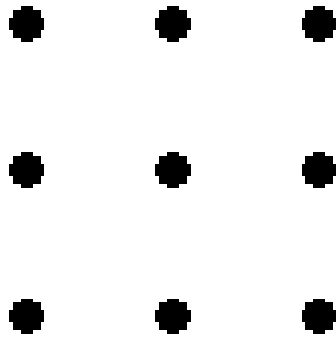


Figure 2: The nine dot problem (Anderson, 1980)

Solution to puzzle

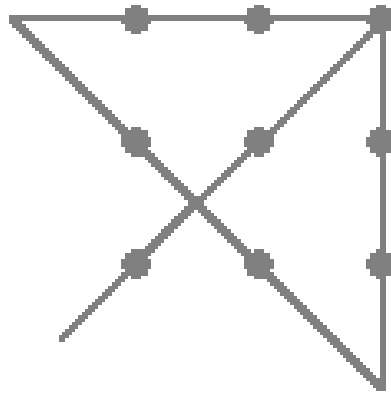


Figure 3: Solution to the nine dot problem (Anderson, 1980)

The answer above demonstrates that a solution to the nine dot problem can only be found when an individual draws lines outside the outer box created by the dots (known as the dot box). It has been thought that the nine dot problem is a good example of an insight problem. Individuals often solve the problem after gaining an insight that they must draw lines outside the dot box. In their experiments Weisberg and Alba (1981) gave some subjects this insight as a clue (that a solution could not be produced unless they drew lines outside of the dot box). They found that subjects who were provided with this clue did not solve the problem immediately after the clue was provided and overall were not much faster at solving the puzzle than subjects not given the clue. Further, after one week they retested the individuals and could find no increased success rate by those individuals who had previously been shown the solution. They conclude that fixation on misleading information (in the dot problem the dot box) and subsequent

insight were not essential for solving the problem. Further they suggested that the fact that individuals did not solve the problem immediately when presented with the clue was an indication that no special process was in effect in solving these problems, and that insight and fixation were terms best un-associated with problem solving.

Now, thirty years on, the analysis of this experiment seems puzzling. Weisberg and Alba set four claims they wanted to test about IBPS:

- 1) The nine dot problem is easy and only made difficult by the misrepresentation of the problem.
- 2) When fixation is broken solutions to the problems should come quickly and easily.
- 3) Solutions to the problems are relatively independent of task specific knowledge.
- 4) Solutions to insight problems will not be forgotten.

They suggest that all four of these assumptions about IBPS are shown to be false by their experiments. Let us leave aside 3. A recent study by Rietzschel et al. (2006) suggests that domain specific knowledge may be important for creative thinking but we may think that being dependent on task specific knowledge is not a central feature of our intuitive understanding of IBPS.

Point 1 aims to show that fixation is not a component of IBPS. Their point is that since, in the cases where clues are given and in the cases where they are not, subjects still require a period of time to work out a solution, fixation alone cannot be responsible for the difficulty of the nine dot problem. This conclusion is backed by the data. But it is aimed at a straw man. IBPS suggests that revelations will be sudden and lead to an insight that facilitates a solution. It does not suggest that insight will immediately provide an answer to a particular insight problem. In the nine dot problem, although it is clear even after the insight that the lines must extend outside the dot box, finding a solution still requires a solution to be worked out utilising the new understanding of the problem. We should therefore not expect that the difficulty of the nine dot problem is entirely based on fixation. This also addresses the second point. Our intuitive understanding of IBPS suggests that insight will come suddenly, not that problem solutions will.

Point 4 aims to undermine a commonly held belief that insight solutions will not easily be forgotten. Early studies by Scheerer (1963) had pointed to this conclusion. Weisberg and Alba (1981) reported that subjects who failed to solve the nine dot problem and who were subsequently shown the solution did not have an increased chance of solving the problem one week later. Weisberg and Alba (1981) took this as suggestive of the fact that insight did not improve recall of problem solutions. But, as pointed out at the time by Dominowski (1981), the individuals tested never actually underwent an episode of insight, with the solutions being presented to unsuccessful problem solvers after a failed attempt at solving the problem. We might think that such an effect is not a necessary component of a special process view. But even if we do hold it as a central tenet of the special process view subsequent studies have suggested contradictory findings to those of Weisberg and Alba (1981).

Similar methodology to that of Weisberg and Alba (1981) was utilized by Dominowski and Buyer (2000). They compared problem solution recall in subjects who solved the insight problems themselves, and in subjects that did not solve the problem but were shown the relevant answers afterwards. They utilised the nine dot problem as Weisberg and Alba (1981) had³⁰. Their study suggested that individuals who solved problems by insight were significantly more likely to solve problems at a second sitting (98% success rate at problem tasks presented a second time) than those who did not solve the puzzle and were informed of the answer after the first failed attempt (61% success rate at problem task presented a second time)³¹. Further, these individuals who actually solved the problem through insight solved the problems quicker than those who had only received the answers rather than solving the problems themselves on the prior occasion. Increased speed and ability to resolve problems at a second sitting is known as the *resolution effect*. Individuals who were unsuccessful on their first attempt and were subsequently shown the answers did show an increased resolution effect, but the effect was significantly less than in individuals who had solved the problem themselves. Both these results contradict Weisberg and Alba's (1981) claim that insight based problem solvers did not show increased speed or recall in insight problems after a second sitting. It seems that Weisberg and Alba (1981) were premature in suggesting that solving problems by insight does not improve recall of problem solutions. No direct conclusion can be drawn from these results and a difference in recall between IBPS and non-IBPS. The results were in agreement with Buyer and Dominowski's (1989) result that showed increased recall on general problem tasks. It does support the idea that solving problems rather than being presented with problem solutions improved recall.

Weisberg and Alba (1981) tested whether the nine dot problem could be solved without insight. Unsurprisingly their studies found that the nine dot problem can be solved without insight. This seems obvious (the answer can be found through chance with a pen and paper). These experiments were seen to support the business as normal view. But this seems hasty. The fact that the nine dot problem can be solved without insight does not indicate that IBPS does not involve a special process. At best it shows us that the nine dot problem should be describe as a *hybrid problem* (Weisberg, 1995), one that can potentially be solved by IBPS or other problem solving means. This highlights a general point about insight problems. The problems are designed to bring about an 'Aha' experience or to be solved by insight. However the categorisation of IBPS is not that it specifically solves these problems. IBPS should be seen as a means of problem solving not as the (only) way in which specific problem tasks can be solved. This early analysis of the nine dot problem therefore can at best be seen as not supporting either

³⁰ They also tested the card problem, the gold dust problem, a version of the water jar problem and four other insight problems.

³¹ Testing across problems indicated that neither group showed an increased ability to problem solve.

the business as normal view or special process view. As we have seen it is helpful in clearing up some misconceptions of IBPS. The following points should be incorporated into our understanding of IBPS:

- 1) Insights arrive suddenly, not solutions.
- 2) IBPS should be seen as a type of cognitive process that is utilised in problem solving, not as a way of categorising particular problems.

2.3.3 Cluster Analysis

An impressive study by Gilhooly and Murphy (2005) aimed at examining whether a range of problems commonly ascribed as insight problems could be separated from non insight based problems. 24 insight based problem tasks and 10 non insight based problem tasks were accessed along with a range of general tasks for measuring cognitive ability, including tests for fluid and crystallized intelligence³². *Fluid intelligence* is a measure of an individual's ability to solve and analyse problems, to think rationally and analyse data. On the other hand *crystallized intelligence* is the ability to utilise prior experience, skill and knowledge in accomplishing tasks. Tests of short term and working memory capacities as well as ideational fluency and flexibility were undertaken by participants. Ideational fluency is the ability to generate a large number of problem solutions, such as providing a large number of common uses for a brick (a classic creativity test included as part of the Torrance tests³³ (1974)) using a brick as a paper weight, or as a door stop, to break a window etc. Idea flexibility is the ability to generate varied and unusual problem solutions, for example using a brick to weigh you down in water.

The results were analysed to see if individuals' ability in solving particular insight or non insight problems was a predictor of their ability in solving other problems within the same type. The results indicated that individual performance that was either above or below average at nearly all of the insight tasks that were tested for was a significant indicator of performance in associated insight tasks. The same was true for non insight tasks. The study also investigated whether distinctive cognitive abilities could be linked to IBPS. IBPS was linked to ideational flexibility; that is an ability to generate varied and unusual solutions to problems correlates with an ability to solve insight tasks. This analysis is a strong indication that IBPS relies on distinct cognitive abilities from those problem solving processes used to solve analytical tasks, because individuals who demonstrated an over average or under average ability to solve individual insight tasks showed an increased likelihood to have the same disposition to all insight tasks and the same for non insight tasks. This suggest that individuals were utilising specific cognitive processes to solve the different task sets.

³² Including spatial insight problems such as the nine dot problem, the triangle problem (see subsection 2.3.4), and the matchsticks problem, and verbal problems including the steel pyramid, X-ray machine, the pound coin problem, the football scores problem, and the marriage problem).

³³ Torrance tests are have been widely used as a measure of creative ability, see chapter 1 section and Torrance (1988).

2.3.4 Feelings of Warmth

A study by Metcalfe and Wiebe (1987) investigated feelings of warmth associated with insight and non insightful problem solving. Feelings of warmth is a term for the subjective measure of how close individuals think they are to solving a problem (like when using the terms warm and cold in guessing games), and whether or not they feel their efforts are being productive or not. Individuals report on feelings of warmth using a visual analogue scale, by marking a paper scale in terms of their confidence of finding the problem solution, every 15 seconds for a maximum of 10 minutes or until a solution was reached. They utilised both verbal and spatial classic insight problems such as the triangle problem (the triangle below points to the top of the page. Move only 3 circles so the triangle points to the bottom of the page).

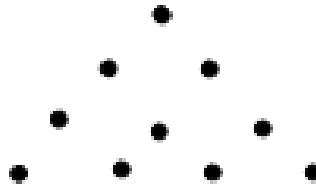


Figure 4: The Triangle Problem

For solution see appendix 1 section 0.5.

Results for the insight problems (12 were used in all) were compared with results from problems that have been shown to be solved incrementally and with algebra problems. They indicated that both in incremental problems and algebra problems feelings of warmth were an indicator of solution imminence. The reported feelings of warmth increased as time passed towards resolution of the problem. This was significantly larger than feelings of warmth associated with insight problems where they tended to increase less incrementally with lower overall feelings of warmth scores but with a rapid increase just before solutions were reached.

This evidence suggests that there is a difference between insight and non insight based problem solving. The feelings of warmth encountered in non-IBPS can be seen as supporting the view that non-IBPS solving proceeds incrementally. Progress towards solving those tasks can be successfully reported by the problem solver. This is in line with the general picture that in NIBPS the perceived way of proceeding is correct and will eventually be fruitful. On the contrary in IBPS problems are not experienced by the problem solver as being solved incrementally and so progress towards a solution cannot be accurately reported on. This is in line with the view that within IBPS no clear solution path is perceived. This might explain some of the surprise involved in the arrival of the solution as an insight.

It might be suggested that feelings of warmth are the only distinct difference between IBPS and NIBPS and that as such they do not show a distinct functional difference between the two processes. First off, it is correct to say that feelings of warmth may not be directly associated with functionally different

cognitive processes. But if the general interpretation is correct, feelings of warmth do have an effect on NIBPS in that they stimulate or function as an encouraging sign for individuals to continue with their current engagement in the problem. They will therefore have a functional dimension as they will feed back into an individual's motivation for continuing or not continuing the current problem solving activity. Secondly, if all problems are solved in the same manner, that is incrementally, then we might ask why some tasks are accompanied by feelings of warmth and others are not.

One explanation is that the problems are of different types; although they produce different feelings of warmth they are nevertheless all solved by incremental process. But this just brings us back to the question of why some incremental changes result in feelings of warmth while others do not. One solution to this problem is to reject the hypotheses that all problems are solved incrementally. Experiments that monitored the heart rate of individuals solving standard problems versus insight problems showed an increased heart rate on standard problems as the time to solution approached (Jausovec & Bakracevic, 1995). These results correlated with feelings of warmth results. This again has been suggested as evidence that standard problems are solved incrementally whereas insight problems are not. It is evidence that feelings of warmth associated with incremental problem resolution are accompanied by physiological changes that are not present in solving insight tasks. The results from feelings of warmth experiments require an explanation. They present a problem for the business as usual view, a problem that is resolved by adopting the special process view.

2.3.5 Verbalisation Experiments

In Schooler et al. (1993) *verbalisation experiments*, subjects were asked to verbalise their thinking as they proceeded with a series of problem solving tasks. The experiments were carried out with insight and non insight based problems: three insight, four non insight problems. The study showed a difference in performance between insight and non insight tasks when verbalisation tasks were carried out. Verbalisation during insight puzzles had an increased negative effect on problem solving performance, compared to non insight problems. The experiments suggest that active verbalisation inhibited the solution of insight tasks but not non insight tasks, and as such were evidence that individuals use different cognitive processing to solve these two types of problem. Further, the experiments have been interpreted as suggestive of the fact that IBPS does normally involves non-verbal processes that are negatively impacted by the activation of verbal processing.

More recent studies however (Fleck and Weisberg, 2004 and Gilhooly et al. 2010) raise important questions about these findings. Gilhooly et al. (2010) tested 16 insight problems (8 verbal and 8 spatial) against 16 non insight (8 spatial and 8 verbal) problems. They found no negative effect of verbalisation on insight vs. non insight problems. They did find that verbalisation had a negative effect on both insight and non insight spatial problems. They suggest that having to verbalise spatial thinking, and in doing so activating verbal processes, hinders the problem solving process when spatial processes were

required. The results are contradictory to Schooler but confirm Weisberg and Fleck's (2004) study that it is spatial problems and not insight problems that are inhibited by verbalisation. In analysing the Schooler et al. experiment, Gilhooly et al. suggest that the effect noted may be due to 2 out of 3 insight problems tested by Schooler et al. being spatial in nature when only 1 out of 4 non insight problems was spatial.

Gilhooly et al. suggest the fact that verbalisation has no effect on solving insight based problems gives support to the business as usual view and does not support the view that un-verbalisable processes are key to IBPS. The argument here is that, like spatial processes, the un-verbalisable process should be negatively impacted by the verbalisation task, since the verbalisation task should interfere with unconscious processing. Since no such impact was noted for insight problems (in general) subjects cannot be relying on non-verbalisable processes to solve these problems.

Key here is whether we should think that verbalisation will have an effect on all non verbalisable processes. Spatial processes may be unique in being impacted by verbalisation. Some studies have shown (see Patrick, 1986 and Sio and Ormerod, 2009) that an incubation effect can be seen even when individuals are asked to perform routine tasks. This may indicate that IBPS can proceed while low level tasks like verbalisation are taking place. The fact that Gilhooly utilised a low level means of reporting is in line with the conclusion that IBPS processes may not have been affected by the need to provide low level reporting.

Overall verbalisation studies show little evidence for a special process view with early evidence to the contrary being most likely due to an effect of verbalisation on spatial tasks. However, little if any support is generated for the business as usual view as no evidence is supplied that suggests verbalisation affects non reportable processes in general. The result that spatial processes are affected by verbalisation provides no evidence for either the special process view or the business as usual view.

2.3.6 The Light Bulb Priming Experiment

Experiments utilising light bulb primes have also shown a distinction between insight based problem solving and other problem solving processes (Slepián et al., 2010). In these experiments participants were asked to complete tasks normally solved by insight (the remote association test-RAT (see subsection 2.4.2 below) and the nine dot problem) and analytical problems (algebraic equations). While the test was underway subjects were primed by the switching on of a light bulb or strip light. In the first experiment the priming was shown to be identifiable with an increased success rate at insight tasks. Further tests ruled out the improved effect being mood related as standard mood ratings showed no correlation with an increased success rate. In the direct comparison between insight tasks and non insight tasks, although both showed an increased effect when the light bulb prime was introduced, the effect was greater in insight based tasks. This general improvement may be due to the light bulb priming for intellectual pursuits in general in the same way as primes that are associated with the elderly cause subjects to walk more slowly (Bargh et al., 1996). The increased effect for insight based tasks over non insight based tasks can be seen as evidence that two

distinct problem solving processes are in operation, with the light bulb inducing IBPS more than analytical problem solving (Slepian et al., 2010).

Again the question can be raised whether the priming results indicate distinct cognitive process or distinct classes of problems to be solved. However if the business as usual view is correct and all processes are solved incrementally then we have to ask why are some problem sets being primed more than others? It could be that the nature of the task has an influence on whether individuals are primed by insight during completion of tasks. For example, maths problems may be more, or less, associated with insight than verbal problems. If all the insight problems used were from topics that had an increased priming effect than those utilised in the non insight tasks then the increased priming effect may not be associated with a distinct cognitive basis for insight but instead be dependent on the type of problem used.

This might be a legitimate concern if the experiments had only utilised the typical set of insight problems. However in the fourth experiment two different types of algebra problems were used. The first group did not require wide scale restructuring to be solved, while the algebra problem utilised to represent insight could be solved by restructuring. When primed with light bulbs, the insight algebra problem had increased solution rates than when primed with other light sources. The algebra problems that were not associated with insight, i.e. did not require large restructuring to solve, were not enhanced by light bulb primes to the same extent as the insight problem. The problems were of the same general topic but some were more complicated than others. It does not seem that there is a basis for considering that the two problem classes would prime insight differently.

This experiment brings the explanation that light bulb primes cued specific problem types rather than the insight based problem solving process, into doubt. Apart from being different in the amount of restructuring required to solve the problems, the problems were of the same type in as much as they were all algebra problems. There is no reason to think that one set of algebra problems will be enhanced more than other algebra problems on the basis that one is more associated with insight than the other. Our original explanation, that the light bulb prime promotes the cognition associated with IBPS and therefore had a larger increased effect on solution rates for insight problem rather than non insight problem, can be seen as the best explanation of the data.

2.3.7 Impaired Frontal Cortex Study

In this study by Reverberi et al. (2005) performance on insight tasks by individuals with impaired frontal cortexes was compared to healthy individuals. The result showed that individuals with impaired frontal cortexes had an increased capacity to solve matchstick arithmetic problems. The matchstick arithmetic problem asks individuals to correct a Roman numeral equation by moving one match (see below for an example). In this study three types of the puzzle were used:

Type A- problems can be solved by moving a matchstick that is part of a numeral to another numeral.

e.g $II = III + I$ (answer: move I from the three to the two, $III = II + I$).

Type B- problems can be solved by moving a matchstick from an equal sign to a minus sign.

e.g $IV = III - I$ (answer: move the match from the = sign to the - sign, $IV - III = I$).

Type C- problems can only be solved by rotating a + sign to make an = sign.

e.g $VI = VI + VI$ (answer: rotate vertical match in + sign to make an equals sign, $VI = VI = VI$).

Type C problems can only be solved by making a tautology.

Type C problems are thought to have an increased likelihood of being solved by insight over type A and B problems. This is because type C problems involve altering the structure of the equation to involve two = signs, whereas type A and B problems can be solved by keeping the basic structure. Tasks that involve underlying restructuring are thought to show an increased likelihood of being solved by insight (Bowden et al., 2005).

Individuals with impaired frontal cortexes showed an increased ability to solve type C problems that are thought to involve increased levels of restructuring, compared to healthy participants. The results can be seen as an indication that the special process view is correct. On the business as normal view where individuals utilise the same processes of solving both insight and non insight problems, it seems unlikely that patients with impaired frontal cortexes should have decreased performance on insight tasks. The fact that they performed differently is an indication that IBPS utilises different cognitive processes than NIBPS.

2.3.8 Neuro-imaging, fMRI, EEG Studies of Insight

Over 20 EEG and fMRI and neuro-imaging studies have investigated the supposed difference between IBPS and NBPS. These studies continue apace, to date there have been many contradictory conclusions (for a discussion on and a recent review see Dietrich and Kanso, 2010). The studies have utilised a wide range of experimental procedures, comparisons have been made between the same individuals solving different problem types and the differences of individuals solving the same problem type in different ways. Within these studies a wide range of classic insight problems such as the candle problem have been used along with anagram problems and the RAT (for details of the RAT see 2.4.2 below). In the candle task individuals are presented with a box containing a book of matches and a box of thumb tacks, and are required to figure out a way of attaching the candle to a wall.) Individuals have been observed utilising EEG, fMRI, ERP and PET scans as methods to observe individuals while problem solving.

Overall imaging data supports the special process view with 19 out of the 20 studies reviewed by Dierich and Kanso (2010) identifying differences in brain activity between individuals solving insight and

non insight based problems. But the exact details of the location or locations of different processes involved in IBPS and the nature of that processing remains unclear.

2.3.8.1 Neurological Differences in Insight and Non-insight Problem Solving

Some early fMRI experiments (Jung-Beeman, 2003, 2004) indicated that insight experiences or ‘Aha’ moments were associated with increased activation of the right hemisphere. They have suggested that the right hemisphere activation, especially within the right superior temporal gyrus or STG may be important for making remote associations. But other studies are contradictory e.g. Qui et al. (2008a) suggesting increased activation of the left hemisphere. The EEG and ERP studies reviewed by Dietrich and Kanso (2010) did not back the observation that the right hemisphere showed increased activation during insight. Overall it seems that there is no specific link between IBPS and the right hemisphere (Dietrich and Kanso, 2010), although there may be a link between insight and the right STG.

The STG is associated with solving insightful problems that require verbal associations and many studies have shown an increased activation in this area in solving verbal insight problems (Jung-Beeman et al., 2004; Kounios et al., 2008; Qui et al., 2006) although for some indication to the contrary see Forstmeier and Rippon (2000) amongst others. Prefrontal areas have also shown increased activation during insight problems (see Lou et al., 2004) although most insight tasks show an increase bilateral activation patterns (see Rose et al., 2002, 2005, and Starchenko, 2003). Further, some studies have shown an increase in activation of the anterior cingulate cortex or ACC (the ACC may be important in detecting cognitive conflict and may be important in the breaking of mindsets (Dietrich and Kanso, 2010). Also, several studies indicate that IBPS is associated with an alpha power decrease (see Kounios et al., 2008, 2006; Sandkuhlr & Bhattachayra, 2008; Danko et al., 2003; Jung-Beeman et al., 2004). I will look more closely at some of these studies when I defend a four stage model of insight.

2.3.9 Conclusions on the Special Process Model

Overall the evidence is highly in favour of the special process view over the business as usual view. Results from cluster analysis, priming studies, impaired frontal cortex patients, and feelings of warmth experiments support this with only verbalisation experiments failing to indicate support for the special process view. EEG, fMRI, ERP, and PET studies all have identified distinct brain activity during insight based tasks and are a strong indication that distinct functioning is taking place. What remains unclear from the gathered evidence is what the functional differences actually are.

2.4 Is a Four Stage Model of IBPS Scientifically and Psychologically Credible?

There is a strong case for considering IBPS as distinct type of problem solving process. In the second half of this chapter I will propose a model of IBPS that is a version of Wallace’s four stage model. Since I will utilise this model throughout the rest of this work in arguing that IBPS is active, in a specific deliberative

sense, within preparation and revelation, it will be important that the model of IBPS proposed is scientifically credible.

In the introduction to this chapter we looked at the intuitive case for adopting a four stage model account of IBPS. The four stage model is quite venerable and has encountered some criticism (see Lubart, 2000-2001). Some of the difficulties include whether the four stage model can be successfully applied to all creative process and whether the four stage structure is too strict to accommodate a fluid interaction of the processes involved in problem solving. (The former problem is solved by considering IBP as one among many creative phenomena). Further problems are associated with the four stage model being based on individual reporting of the creative experience.

As much of the intuition concerning IBPS is based on individual reports of the phenomena, in order to provide a proper scientific grounding I will consider, in turn, whether each of the stages of the four stage model is supported by the psychological and scientific evidence. This will provide a defence against the claim that the four stage conception of IBPS does not accurately represent the process, either because the phenomenology on which the reports are based is epiphenomenal or because the process commonly leads to misleading, after the fact, reconstructions of the experience.

2.4.1 Preparation

Within preparation individuals are engaged in deliberating about the problem at hand, giving up, carrying out further research and re-engaging in deliberation. (This can be seen as the initial stages of IBPS where individuals create and exhaust a search space).

Mumford et al. (1991) break this stage down into several parts, including *problem construction*, *information gathering* and *concept selection*. Problem construction³⁴ consists in the problem solver coming to conceive of the problem at hand in a specific way. In doing so the problem solver brings to the surface key elements of the problem to be solved. In information gathering, the problem solver gathers new information that is relevant to the current question. In concept selection the gathered information is sorted according to specific key concepts that an individual thinks most relevant to the problem at hand³⁵.

Experimental investigation into these processes necessarily involves asking the subjects to purposefully engage in problem constructing, information gathering or concept selection. Here the activity of the subjects can be seen as their intentional engagement in the particular process.

Experiments that directly prompt intentional problem construction show a correlation between engaging in problem construction and obtaining increased problem solving results. Studies include Okuda

³⁴ Often called problem finding.

³⁵ These activities of preparation are not exclusive to IBPS they will be utilised in NIBPS problem solving. The difference between NIBPS and IBPS is that in the former the activities of preparation will result directly in the problem being solved. While in IBPS no solution will be forth coming.

et al. (1991), Rieter, Palmon et al. (1997), and Mumford et al. (1997). In Scott et al. (2004) meta-analysis of problem construction showed that it had positive effect within problem solving (Scott et al., 2004).

The information gathering experiments that have been carried out ask the subjects to intentionally gather more information after a distinct problem construction phase has been completed. Such work leads to increased problem results when the information gathered was relevant to the question at hand (Mumford, 1996). For evidence that concept selection has a positive effect within problem solving, see Mumford et al. (1996) for an individual study and Scott et al. (2004) for a meta-analysis.

The method of actively³⁶ inducing problem construction, information gathering and concept selection presents a problem for interpreting the results of these studies. It would be preferable to have evidence that these processes have a beneficial effect on problem solving without actively instructing individuals to participate in those processes, but without such instructions it seems difficult to devise an experiment within which the effect of such processes can be individually verified. Further, the evidence gathered here only supports the claim that these processes benefit problem solving, not that they are naturally present within problem solving. However these processes are intuitively the type of activities we engage in when problem solving and are commonly reported in problem solving studies (see Getzels and Csikszentmihalyi, 1976) and so the above evidence supports the general proposition that these activities take place within problem solving. Evidence of increased problem solving success when problem finding, information gathering and concept selection are intentionally pursued suggests that they are likely to increase success within problem solving generally even without positive instructions to intentionally engage in those processes.

Overall these three processes provide support for an intuitive way of understanding preparation as the initial stage of IBPS. We can think of these three processes as being interrelated, with individuals returning to different activities in a fluent manner (Lubart, 2000-2001), defining problems more tightly as information is gathered and sorted. The studies highlighted show direct support that the activities of problem construction, information gathering and concept selection increase problem solving ability, but as discussed above this can be interpreted only as indirect evidence that such processes actually occur in problem solving. However, it seems in keeping with this analysis and with the overall picture presented in the problem solving literature that some sort of preparation is required for problem solving.

2.4.2 Incubation

In the four stage model incubation is assumed to occur because the product seems to arrive suddenly and surprisingly within revelation. Incubation represents the period of non conscious restructuring that takes place on the special process view. What evidence is there that this period of ‘not intentionally pursuing’ the question at hand has a positive effect on problem solving? A meta-analysis of 29 experiments on

³⁶ The type of activity that individuals are involved in here is intentional and should be held separately from the deliberative activity I hope to demonstrate is present within preparation and revelation. The mixed nature of the activity within preparation will be examined later.

incubation has been carried out by Sio and Ormerod (2009). All the experiments followed the same general structure, although they incorporated a range of experimental variables. Subjects were set a particular problem (the nature of the problem varied between experiments); a common example was the *Remote Associates Test* (RAT)³⁷. In the RAT subjects are given a series of three word sets and asked to find the word that connects all three (Sio and Ormerod, 2009). For example, the puzzle ‘Walker, Carriage, Bath’ is solved by the word ‘Baby’. In the control test, subjects worked on the problems continuously for a period of time. In test subjects, a break period is introduced, which varied from a few seconds to over twenty minutes, after which they completed the work period. The meta-analysis of the 29 studies showed a significant positive *incubation effect*, in that individuals who underwent an incubation period produced better overall scores³⁸.

Two general hypotheses, the *conscious* and *unconscious work hypotheses*, have been proposed³⁹ to account for the incubation effect. The conscious work hypothesis suggests that the incubation effect is not due to unconscious cognitive work during incubation but is rather due to the incubation period causing an improvement in future conscious processing (Posner, 1973). The conscious work model can be seen as essentially the same as the business as usual model. Explanations of such effects include covert conscious problem solving, where individuals take part in conscious processing of the problem during the break period, and a natural reduction in fatigue due to the break away from the problem. Contrarily the unconscious work hypothesis suggests that problem solving processes are, unconsciously, taking place within incubation itself; this can be seen as essentially the same as the special process model (Yaniv & Meyer, 1987)⁴⁰.

Distinguishing between these two hypotheses was a central aim of Sio and Ormerod’s (2009) meta-analysis. Seven moderators were chosen for analysis: the activity of subjects during incubation, length of the incubation period, length of the preparation period, the ratio of the preparation period to the incubation period, the nature of the problem to be solved and the presence of both solution relevant and misleading cues. The results showed an overall positive effect for incubation, with longer periods of preparation and incubation both producing larger positive effects. Linguistic puzzles, including the RAT, showed a smaller effect than divergent thinking tasks. Concerning the activity during incubation, an increase in positive effect was recorded when basic tasks were used over complex tasks and an increased effect from basic tasks over resting. Sio and Ormerod’s (2009) analysis suggested that this last result can be seen as evidence against the conscious work model. If the incubation effect was due to reduced fatigue, then we would expect resting to produce a larger incubation effect than taking part in basic tasks. Sio and

³⁷ The incubation effect was less in the RAT than other experiments. This maybe because RAT like problems are sometimes solved by insight and sometimes analytically.

³⁸ Assessment of scores was sometimes straightforward (as in the RAT) and sometimes required judges to gauge the relative novelty of the results.

³⁹ Assessment of scores was sometimes straightforward (as in the RAT) and sometimes required judges to gauge the relative novelty of the results.

⁴⁰ Consciousness is seen here to be linked with reportable, high level processing of a problem

Ormerod's meta-analysis showed that this was not the case. The fact that engaging in basic tasks over resting was beneficial for incubation is also evidence against the covert conscious work hypothesis. For it seems that resting would provide more time to engage in such activity, over the time available in engaging in simple tasks. The finding that taking part in complex rather than simple tasks during incubation is less effective, might suggest that complex tasks do not leave enough cognitive space for unconscious processing to take place (Sio and Omerod, 2009). The evidence above supports the conclusion that incubation has a beneficial effect on problem solving and points to an unconscious work model of incubation that is in line with our conclusion above that IBPS represents a distinct type of problem solving process.

Now it has been empirically supported the unconscious work model can provide a useful way of distinguishing between the stages of preparation and incubation. Within preparation the problem is being actively pursued, whereas in incubation, the processing of the problem has no active element whatsoever. Instead further unconscious processes generate the problem solution. The meta-analysis above supports the unconscious work model but gives relatively little indication of which unconscious processes will in fact bring about problem solutions. The actual process involved is a matter of great speculation within the literature. As we noted above restructuring has often be seen as key to IBPS. (No defence of a particular model will be provided here⁴¹. All that is required is that incubation be seen as distinct from preparation⁴²). IBPS should be viewed as containing a period of incubation.

It can be noted that an incubation effect does not provide evidence for revelation. Unconscious processing could be advantageous without leading to an 'Aha' moment. However a number of results pointed to incubation being effective specifically in IBPS as opposed to problem solving in general. Longer periods of preparation (up to 20 minutes) before the introduction of incubation showed a stronger incubation effect (Sio and Ormerod, 2009). One explanation of this increased effect is that in order for incubation to be effective individuals require to have reached an impasse in their deliberations (Sio and Ormerod, 2009). Before an impasse is reached, incubation will have little effect. This is supported by the analysis of experiments where different types of puzzle were used. Here, problems that were less likely to be solved overall showed an increased incubation effect over simpler problems (Sio and Ormerod, 2009). This fits the analysis above as individuals engaged in difficult problems were more likely to reach an impasse and so more likely to gain a benefit from incubation. This gives some support for incubation being more beneficial for IBPS than other forms of problem solving.

2.4.2.1 Evidence for Restructuring in Hindsight Bias Experiments

⁴¹ For a model that fits with the four stage conception of IBPS see Thagard and Stewart (2010).

⁴² Although the unconscious work model is supported by the meta-analysis (Sio and Omeros, 2009) it is not essential for my account. The evidence presented here can be seen as supporting the general conclusion that an inactive period, between periods of active pursuit of a problem, as carried out in the above experiments, has a beneficial effect on problem solving.

Hindsight bias is the phenomenon related to the making of a judgment about whether one could resolve a particular problem. When individuals are in possession of facts that pertain to a given solution they are more likely to suggest that they could have solved the problem. For example in the nine dot problem individuals who are provided with the answer to the problem, but were asked to ignore that knowledge, when making a judgment about their likelihood of solving the puzzle, tend to overestimate their chances of solving the puzzle. It is suggested that hindsight bias occurs because individuals in possession of outcome knowledge might automatically restructure their thinking leaving them unable to access their original representation of a problem (Ash and Wiley, 2006).

Ash and Wiley (2006) suggest that processes that utilise more restructuring, as has been proposed for IBPS, will also have an increased hindsight bias. On this hypothesis they predict that individuals that solve insight problems should show an increased hindsight bias on key components of those problems that lead to restructuring, over the hindsight bias indicated for the key components of the puzzle that do not lead to problem solutions. Ash and Wiley selected two algebra and two insight problems. They included the cups insight problem and the triangle insight problem.

In the cup problem, six cups are in a line. The first three contain liquid. The question to be answered is: describe how you could make it the case, by moving only one glass, that no two glasses containing liquid are next to each other, while keeping three of the six glasses full.



Figure 5: The cup Problem

For solution see appendix 1 (section 0.6).

Ash and Wiley state that B and E are key components of the problem. Individuals who solve the puzzle will have increased hindsight bias for importance levels on these components when compared with the importance scores for non problem solvers even when non problem solvers are presented with the answer.

For both the triangle problem and the cups problem, importance ratings and hindsight bias increased for key components in individuals who solved the insight problems compared with individuals that did not, although the level was only significant in the triangle problem. Ash and Wiley suggest that this increase in importance ratings is an indicator that increased restructuring is taking place in insight tasks as opposed to algebra tasks. These results are largely supportive of the special processes view. Further they are an indication that restructuring is important in IBPS.

The folk conception of IBPS is entirely silent about what sort of process is producing the incubation effect. The meta-analysis above supports the unconscious work model but gives relatively little indication about what unconscious processes will in fact bring about problem solutions. I will not venture to discuss the nature of the processing within incubation; for some recent indicators of the events of incubation see Stokes (2007) and Thagard and Stewart (2011).

2.4.3 Revelation

Revelation is the ‘Aha’ experience, the moment of realisation or understanding that seems to come out the blue and as such is central to both the four stage model and IBPS. This phenomenon of realisation is one reason why this type of problem solving is so puzzling. Topolinski and Reber (2010) separate four main characteristics of the ‘Aha’ experience that fit with the intuitive understanding of the phenomenon:

- 1) **Suddenness:** The problem solution comes to the mind suddenly or abruptly, that can perhaps best be described as popping up.
- 2) **Ease:** Working on the problem after the sudden revelation is often easy and straightforward.
- 3) **Positive affect:** The ‘Aha’ experience is connected to a positive affect, a feeling of pleasure, satisfaction or elation that is wrapped up with the product.
- 4) **Truth and Confidence:** The insight has an increased likelihood of being judged as true even before further evaluation has taken place.

Recent experiments by Kounios and Beeman (2009) have tried to measure the ‘Aha’ experience using EEG and fMRI experiments. In these experiments individuals were given RAT like problems to solve while being monitored by EEG and fMRI machines. The RAT was chosen because subjects under experimentation show two distinct behaviours in solving RAT puzzles. Some puzzles are solved analytically, by participants running through a variety of possible word clues, while other problems are reported to be solved by insight. The collected EEG and fMRI data was correlated with individuals’ reports on whether they solved the problem through insight or analytical methods.

These experiments provide two indicators of support for revelation. Firstly individual reports were taken under experimental conditions. Reports under such conditions are likely to be more accurate than autobiographical reports made after the fact, as participants are asked to report truthfully within a short time frame. This removes one reason to be suspicious of such reports and thus helps support the commonly held assumption that the revelation experience takes place. Secondly, comparing the EEG results from episodes where individuals reported an ‘Aha’ experience to episodes where individuals reported solving the problem analytically showed different brain activity. Specifically, EEG results showed that solutions solved by insight included a burst of high frequency activity occurring at 40-Hertz for 300ml seconds prior to the test subject pressing a button to indicate that a problem solution had been

obtained. This suggested that subsequent reports of the 'Aha' experience were not just different post hoc descriptions of the phenomenology but were underpinned by a specific cognitive episode that was detectable with EEG. This provides some evidence that the 'Aha' effect is an indicator of the specific type of functioning associated with IBPS. There is no evidence here that the conscious nature of the 'Aha' experience is playing a distinct functional role. The experience may be epiphenomenal but its correlation with specific brain activity can still be an indication that the 'Aha' experience is related to a specific type of problem solving process.

2.4.4 Verification

In verification the product revealed within revelation is thought to be assimilated into our current understanding and made available for further scrutiny (Wallace, 1926). As has been shown by Weisberg and Alba (1981, see above) in classic insight puzzles such as the nine dot puzzle revealing the clue does not facilitate individuals immediately solving the puzzle. Likewise insight brought about in revelation does not always provide the answer to a problem. But insights often do eventually lead to task resolution. With the insight in hand an individual enters a new search space within which a solution can be found. Verification then can be seen to be much the same as preparation with some of the same processes being in effect. Unlike preparation there is little need to frame the problem as the new revelation has already provided a new search space. Verification is likely to have two outcomes: either a solution to the problem is found or the subject again reaches a state of impasse where the current search space is exhausted without result. Therefore verification can on some occasions be seen as important for initiating further problem solving activities beginning again with preparation. It is possible that many rounds of the process may be required before a solution to a problem is reached.

2.4.5 Constituents of Insight Based Problem Solving

The evidence presented above provides us with strong support for adopting an account of IBPS that follows the four stage model. We have taken the following elements as central to our account of IBPS:

- 1) Some form of the special process view is correct.
- 2) Processes within preparation affect downstream problem solving.
- 3) Preparation can be distinguished from incubation.
- 4) Incubation takes place and has a positive effect on problem solving.
- 5) The 'Aha' experience takes place and is useful in indicating a specific type of problem solving.

To summarise, we have substantial evidence from neural imaging, feelings of warmth experiments, and cluster analysis that IBPS is a psychologically distinct form of problem solving process. This upholds the

special process view. We have supporting data from priming experiments and the impaired frontal cortex study. This is also supported by the meta-analysis of incubation. Further the hindsight bias experiments give us some reason to suspect that restructuring plays a role in incubation.

With regard to understanding IBPS in terms of the four stage model, we have evidence that intentionally pursuing problem construction, information gathering, and concept selection within preparation increases success rates with problem solving tasks. This indirectly supports the claim that these activities are taking place within and are beneficial for IBPS. Importantly, we have good scientific evidence that incubation and revelation take place and are associated with IBPS. Further we have good evidence that incubation has a positive effect on problem solving and that it is specifically useful for IBPS. We have specific data that suggests that revelation is a genuine phenomenon that can be correlated with distinct neural activity. Now that a defence of IBPS as a scientifically robust phenomenon is in place I can, in chapter 5, provide an account of how individuals within IBPS are active. But before that I will examine whether IBPS conceived as a four stage model can be understood in terms of intentional mental action.

Chapter 3

Mental Agency and the Bounds of Action

3.0 Introduction

In recent years there has been a growing debate about whether mental events are agentive. A variety of mental events, such as imaginings, the recalling of memories, deciding, etc., have been discussed⁴³. One way in which philosophers have tried to answer this question is to describe mental events as intentional actions (Strawson 2003; Mele 1997, 2010). Intentional actions have a long philosophical tradition from Aristotle (2009) to Anscombe (1957). In trying to make sense of the wide variety of human and animal behaviour in our lives we make a division between things that we do and things that happen to us (Wilson and Shpall, 2012). I go hunting, the lion attacks me. The things that we do, in a primitive way, we have often called actions. Understanding the nature of action helps us make this separation. Characterising behaviour in this way allows us to navigate and understand our complicated physical and social environment, since it provides us with evidence for judging the effect of our and our fellow's actions on each other. Without being able to make this distinction, it is difficult to understand what it would mean for an event to be down to me. It is this central place in our understanding of ourselves and of our society that gives the debate about what is, or is not, an intentional action its import. Because of this position intentional actions have been held up as paradigm exemplars of agency (Wilson and Shpall, 2012). That is, they can be said to be down to us, or things that we do, in the strongest sense of these phrases. Endorsing a behaviour as an intentional action, therefore, has consequences. It links that behaviour to many basic psychological concepts such as motivation, pleasure, guilt, esteem, praise, blame, etc., and with social or legal concepts such as status, ownership and responsibility. For it suggests to us that we are the originator of such behaviour and, as such, appropriate candidates for such attitudes.

If mental events are intentional actions then they can be seen as agentive in the same way, to the same extent, and entail the same consequences, as physical intentional actions. But there is a question to be asked here. Is intentional action the correct way to characterise the agentive nature of mental events? Attempting to answer this question is one goal of this chapter. In investigating this topic, I will utilise several different accounts of intentional action. By introducing increasingly liberal accounts of intentional action we will be able to describe an increasingly large percentage of mental activities as intentional actions. The type of mental activity that can be properly described as intentional actions will vary depending on the details of the particular accounts. By progressing in this way, it is hoped that the particular difficulties involved in describing mental events as intentional actions will become evident.

It is the claim of this thesis that IBPS can inform our understanding of mental agency in a heretofore unexplored way. The point of the investigation into the extent to which mental activity can be understood as mental action is to inform our agentive understanding of the process of IBPS. Uncovering

⁴³ For a noteworthy collection see O'Brien and Soteriou (2009).

the full nature of the agency involved in IBPS will take us beyond the confines of this chapter. But the first step is to examine the extent to which mental events including IBPS can be described as intentional actions. If IBPS can be understood as an intentional action then we would have answered the concern that creativity cannot be agentic because the creative phenomenon of IBPS is involuntary.

As I have suggested in chapter one IBPS is a four-stage process. Specifically I argued for the following points:

- 1) Some form of the special process view is correct.
- 2) Processes within preparation affect downstream problem solving.
- 3) Preparation can be distinguished from incubation.
- 4) Incubation takes place and has a positive effect on problem solving.
- 5) The 'Aha' experience takes place and is useful in indicating a specific type of problem solving.

With these points in mind we put forward the following view of IBPS (section 2.4). In preparation we actively engage in processes like problem finding, information gathering and rational deliberation. Preparation is followed by incubation, where on the special process view, distinctive processing takes place and brings about revelation. Revelation is the 'Aha' experience of coming to a new insight about the problem at hand. This new insight is then utilised in verification and an answer to the problem at hand is produced.

To set the scene of my investigation into mental activity as mental action I will begin by introducing Davidson's causal account of intentions (1980a). I then discuss some basic difficulties with causal accounts before setting out the two conditions of *the simple view* of intentional action. I then examine Galen Strawson's (2003) hard-line approach to mental action. Even on Strawson's (2003) sceptical account, several types of mental activity, such as *priming* and *shepherding* actions, will be intentional. I then proceed to a less constrictive interpretation of the simple view. Here we will encounter some of the major difficulties of understanding mental events as intentional action. One of these difficulties is the *access puzzle*. The access puzzle arises because of the structure of intentional action. On a causal account of intentional action for an action to be intentional the result of the action must be represented by the content of the agent's intention.⁴⁴ But in many types of mental event the content of an intention cannot be linked with the result of the mental activity because the agent cannot represent the action result within their original intention. The access puzzle is particularly problematic for IBPS, because in IBPS the goal of the process is to answer a particular problem, and this will require the bringing about of new content.

⁴⁴ For non-causal accounts the access problem may not be a difficulty. But non-causal accounts have other difficulties see subsection 3.1.1 below. Within this thesis I will only attend to causal accounts of intentional action.

We will then look beyond the simple view. The simple view has been criticised as being too demanding, in that it fails to accommodate some types of physical activity we normally think of as intentional actions. This has motivated the adoption of more complex accounts of action such as Mele and Moser's (1994) *advanced causal account* (ACA) and Mele's *general casual theory* (GCT) that is an attempt to apply ACA to mental events. It seems appropriate, then, to examine the extent to which these advanced causal views can accommodate mental events as intentional actions. Under GCT more mental events can be described as intentional actions than under the simple view. But describing mental events as intentional actions under GCT will again be limited by the access puzzle, and as a result no satisfactory account of intentional action is available for IBPS. This examination of these causal views, ACA and GCT, uncovers some telling aspects of particular mental processes that increase the difficulty of describing them as mental actions. These difficulties reflect the difference between mental and physical activity. Although such difficulties can be accommodated for some mental events, it can be suggested that the difficulties are an indication that characterising mental events as intentional actions may not fully capture the essence of what makes mental events in general agentic.

One of the main reasons for introducing these different accounts is to suggest the difficulties associated with understanding whether these accounts are applicable to IBPS. Progressing through these different accounts may seem time-consuming and at times unmotivated, but I hope by the end to have provided a comprehensive assessment of the extent to which causal accounts of intentional actions could potentially accommodate IBPS. In making such an assessment I want to be as inclusive as possible, so even if some accounts are problematic, understanding the reasons why IBPS cannot be successfully understood as an intentional action on each account will justify us looking further afield. In chapter 4 we will examine whether accounts of responsibility for action consequences can be utilised to produce an account of responsibility for the products of IBPS. In the fifth chapter I will look beyond accounts of agency based on intentional actions and their consequences. Let's begin looking at intentional action by introducing intentions.

3.1 Davidson: The Nature of Intentions

We can begin the discussion about whether intentional action is a good model for mental agency by getting to grips with intentions themselves. The Davidsonian model (1963) of intentions suggests that the attribution of *pro-attitude*–belief pairs to actions is a necessary condition for intentions, where pro-attitudes are attitudes that are directed towards an action. Pro-attitudes cover a range of mental states. The classic partner for a belief is a desire, but Davidson suggests a range of mental states, including wantings, urgings, moral views, aesthetic principles and social norms⁴⁵. So an example of a *belief–desire* pair that forms an intention would be a belief that tomorrow elephants can be seen at the zoo and a desire to see

⁴⁵ For ease I will refer to pro-attitude–belief pairs as belief–desire pairs; see Davidson (1963) for a full account.

elephants at the zoo tomorrow. The belief–desire pair is playing two distinct roles in Davidson’s account. It is providing the motivation for the intention, and it is providing a rationale by which the intention can be made understandable, both by the agent and by observers (Davidson, 1980a). This belief–desire pair is an intention, and the first condition for a resulting action being intentional (i.e. your going to the zoo and seeing an elephant) is that you have a reason for so acting (in this case the desire to see elephants, and the belief that there are elephants in the zoo) (Davidson, 1963). For Davidson, further factors have to be included before we can arrive at a full account of intentional action: it is necessary that there is a causal relation between your reason, in the form of a belief–desire pair, and the bringing about of your intention (Davidson, 1963a). According to Davidson, then, what can be said to be an intention that forms a part of an intentional action is a belief–desire pair that plays the right kind of causal role in the bringing about of what you intend.⁴⁶

It is worth our time to explore briefly in what way Davidsonian-type intentions interact with other mental states and the world at large to function effectively. For example, your intention to visit the zoo this afternoon is not an all-out commitment to going to the zoo this afternoon. Only some goings will be desirable. Just because you intend to go to the zoo, we do not want to say that you are intending to go under any circumstance, but only make the narrower claim that you intend to go given how things stand presently and perhaps foreseeably⁴⁷. This indicates that belief–desire pair states must be in a position to interact with your overall web of beliefs about the world. We might consider that an intention is under the constraint of a ‘current state of play’ condition. If conditions are right, the belief–desire pair will be causally effective in bringing about what you intend and this will be bound by a set of conditions formed (most probably implicitly) in the formation of the original intention. We have now a view about what intentions are: they are reasons for action embodied in the form of belief–desire pairs. We need to look briefly at the question of how and if such reasons take up a causal role in a physical description of action. This will give us a better understanding of a general underlying metaphysical characteristic of causal theories of action, that reasons are causes. This characteristic of causal theories will be important for ruling out accounts of IBPS as intentional actions.

3.1.1 Intentions as Reasons and Reasons as Causes

⁴⁶ In Davidson’s (1980b) account a further overall judgement maybe required. He suggests that the formation of an intention to act may require that the intention is formed through an overall judgment about your pro-attitude and your belief about the word. If you intend to wash the car your intention arises through your all-things-considered judgement that washing the car is something you want under the conditions of what you currently believe about the world. The inclusion of this judgement in the formation of intentions is intended to solve a difficulty in agents holding intentions that are not followed by actions. Another objection raised by Mele (1997) provides a counterexample where this rationalising aspect seems to separate from a basic account of intentional actions: that, for example, you can choose to sing a song you wrote just because you possess a desire to sing that song.

⁴⁷ For discussion on this point see Bratman (1999).

In Davidson's view, intentions are reasons and reasons are causes (Davidson, 1969). That is, reasons can be seen as the causes of actions. Accounts of action that hold this view have been termed 'causal accounts'. There is a competing view that reasons are solely explanations of actions and should not be considered causes. The difficulty with such a view is that it seems to give no explanation of how it is that we are motivated by our reasons. If reasons play no motivating role then why are they relevant to understanding our behaviour? A rational explanation alone provides no justification for why we should consider that we were moved to act by the reasons that form that explanation. But if reasons play a causal role and we can understand them as relevant causes for our behaviour then we have a justification for seeing the behaviour as motivated by the agent's own rationality.

Davidson (1980a) offers the following challenge to the non-causalist. Sometimes situations arise in which two rationalisations of an action are possible. I can, for instance, make a remark that amuses one of my fellows, Tillmann, and infuriates another, Dave. Further, I can be in possession of separate reasons for wanting to do the two things. If these reasons are not causes, how can it be that one set of reasons explains my behaviour over the other? According to Davidson, the reason that actually does the causal work should be seen as the reason that rationalises the action. Mele (1992a) suggests no successful solution to this problem has been offered by the non-causalist. The best attempt to answer the challenge comes from Wilson (1989)⁴⁸.

Wilson (1989) tries to meet Davidson's challenge by building in an agent's reason for acting into their intention. If I alter the example above to capture Wilson's account, I would make a remark with the intention that had a specific reference to either, irritating Dave, or alternatively of amusing Tillmann. I could not be in a position where I just intend to make a remark (or other action) without indicating in my intention what I want to achieve by so acting. Thus the nature of the intention would dictate what explanation rationalised my action.

In reply, Mele (1992a) introduces an example where an individual holds two intentions, both of which could rationalise the action, but only one of which is required to rationalise the action. For instance, an individual might have intentions to open the window to freshen the air and to open the window to announce that people on the bus are smoking. Both are explicit intentions of the nature Wilson suggests solves the puzzle but even so when two such specific intentions are present it is not possible to know which of the intentions should be considered pertinent. Without a causal connection between the intentions there is no way of separating which intention, and therefore which set of reasons, we should consider relevant to the action. Further Mele (1992a) objects that Wilson's account makes intentions too complicated. In some cases our reasons for acting are complicated, and cover a wide range of motivations,

⁴⁸ A further answer to Davidson's challenge comes from Ginet (1990). Mele's (1992) answer to this challenge can also be found in Springs of Action.

considerations and concerns, but it seems unnatural to think that an individuals' intentions must always be couched so as to include their complicated motivating considerations⁴⁹.

In this thesis I will side with Davidson (1980a), Mele (1992a), Velleman (1992) and Bratman (1984) in thinking that the best explanation of actions is a causal one. I will not argue for the basic contention of causal theories that intentions, as reasons, are causes of actions. Nor will I examine in detail whether mental events can be described as intentional action by non-casual accounts of intentional action⁵⁰. I will not argue for any particular causal view (due to the constraints of space and focus) because my argument against understanding IBPS as a type of intentional action will depend on a basic feature held in common by causal theories of action: that intentions are mental states, and must play a causal role in order for an action to be intentional. My argument against understanding IBPS as an intentional action will depend on seeing actions as having these features.

I will adopt the approach of examining a range of views that are typical of the debate concerning the nature of causal intentional actions. Presenting the different causal accounts of intentional action also provides a way to examine the nature of mental events. In keeping an open mind concerning the definitive nature of causal account of intentional action, it is hoped that the general problems of understanding mental events as actions can be uncovered. In any case, drawing general conclusions about the nature of mental activity will not be dependent on any particular causal view, and this provides another justification for examining several such views. With the depth of my commitment to causal accounts of intentional action outlined, we can move on and remove another complication with accounts of intentional action: the different ways actions can be individuated.

3.1.2 Two Ways of Individuating Actions

You are sitting on a park bench with your nephew when it begins to rain. He is crying at the loss of his boat in the pond and is in need of cheering up. You open up your umbrella in a comical fashion, bring it to

⁴⁹ Consider a game of basketball, I have the ball and want to pass to Dave because I am being double teamed and because Dave is a great shot. Through some sleight of hand on my part I open a small gap and pass to Dave. It does not seem that I should be required to specify the exact reason for me wanting to pass to Dave in my intention for this activity to be considered an action. If we accept this case as an action, then we are again in a position where we will have no way of distinguishing which set of reasons motivated it, those associated with me being double teamed or the unequivocal fact that Dave is a great shot.

⁵⁰ It can be asked whether I am justified in only examining causal accounts of action. IBPS is a phenomenon with a particular phenomenology that is quite distinct from an action. Theories of actions that depend on a common phenomenology between mental events in their ascription of actions are therefore unlikely to provide a useful way of interpreting IBPS as an action. Here time and space spent in this chapter is an indication of the space that would be required to accommodate a comprehensive investigation of whether IBPS can be understood as an intentional action under all type of action accounts. I will then 'pin my colours to the causal mast'; in doing so I am in good company (Davidson, 1980a; Mele, 1992; Velleman, 1992 and Bratman, 1984).

your shoulder and in so doing you scare the ducks in the pond. Your behaviour can be described in many ways, your moving your hand to open the umbrella, your opening the umbrella, your sheltering from the rain, your amusing your nephew, or even your scaring the ducks. The question of individuating action is concerned with how we can successfully divide up this behaviour. Should we see it as one action or more? That is, should we see these actions (you might say one action) as identical? The question is highlighted by Anscombe (1957), but the two main views are championed by Davidson (1969) and Goldman (1971). Davidson suggests that events are identical if they have the same causes and effects. This would suggest that there is only one action in the above case; that your moving your fingers, opening the umbrella, and amusing your nephew are all descriptions that refer to a single action, because they are the same event, and this is the case because they have the same causes and effects. All the above statements are descriptions that refer to the same event on the basis of them having the same causes and effects. The claim is then that the descriptions, 'my moving my fingers' and 'my opening the umbrella', describe the same event only if they have the same causes and effects⁵¹. Davidson suggests that even if they all refer to a single identical action, 'they all are caused by the intention to open the umbrella', only some of the above descriptions are intentional because only some of them refer to parts of the action that are rationalized by specific belief–desire pairs. For instance, your belief desire-pair to amuse your nephew, and believing that opening the umbrella will amuse your nephew, rationalises the description of the action as opening the umbrella to amuse your nephew, and makes that action intentional under that description; it doesn't rationalise the action of scaring the ducks although it is still a cause of that event.

Goldman, on the other hand, argues that all these descriptions pick out separate actions. He is moved by the fact that the descriptions seem to have a specific order; that trying to move your fingers, moving your fingers and opening the umbrella follow one after the other. He suggests there is a difference between the relationships of these actions: the later actions are dependent on the earlier action while the former ones are not dependent on the later. Further, the agent seems to do the later actions by doing the former ones, and it is not true that they do the former by doing the later, e.g. move the fingers by opening the umbrella. This asymmetry of relation suggests that these actions are not identical and so cannot be the same action. Therefore the above statement refers to a number of different actions. Goldman's account is not without its own difficulties. For in suggesting that all these descriptions are individual actions, we multiply the number of actions we take part in by the number of different descriptions that are appropriate. So in the above example, you will perform several actions, including moving your fingers, opening the umbrella and amusing your nephew. This multiplies the number of actions that we do by the number of available descriptions, which will normally be quite large. A better explanation — at least a more

⁵¹ We can leave aside any debate about whether we should think events should be individuated on the basis of cause and effect or whether the above statements actually do refer to the same events.

parsimonious one — is to understand this as the performance of a single action of opening your umbrella to amuse your nephew⁵².

It is not important for our present purposes to decide between the different merits of these accounts. Both methods of individuating actions will be acceptable. We need to be clear that the problem of individuating actions is separate from the access puzzle (section 1.0). In both problems the way in which you describe an action is problematic. Different descriptions are applicable to a specific behaviour, and only some of those descriptions are going to turn out to be descriptions of intentional actions. The access problem arises because of the need to match the descriptions of action results with the content of our intention. Only some of these descriptions will refer to intentional actions. But this is not a matter — as in the case of individuating actions — about the metaphysical nature of actions; rather, it concerns what descriptions can be seen as intentional action given the content of the intention. In accounts of intentional action the content of the intention will matter for discerning what actions are intentional not whether a particular behaviour is best understood as a single or as multiple actions. In understanding the metaphysical nature of action the content of an intention does not seem to play a role separate from it being understood as a cause. But here although the intention is a cause of the behaviour it can be seen as a cause of all the behaviour not just the behaviour represented in the intention and so that representation is not playing a role in deciding what is or is not an appropriate individuation of action.

3.2 The Simple View

With a notion of intentions and a clarification of the lines and limits of our enquiry in place, we can turn to the examination of mental agency via an examination of increasingly permissive accounts of intentional action. We will begin with the stringent requirements of the simple view. In one version of the simple view, for an action to be intentional it must conform to two conditions. Firstly, for an event to be an action an individual must have a specific intention to bring about the result of her action. I will call this the *condition of tight fit*. Secondly, there must be a causal relation between the original intention and the bringing about of the action result (Adams, 1986).

The condition of tight fit amounts to the requirement that for an event to be an intentional action, the individual concerned must have formed an intention to bring about that specific event. That is, forming an intention to go to the post office is a necessary condition for describing our going to the post office as an intentional action. Our action must fit the content of our intention (at least under one description of that action (Davidson, 1963)). For the moment we can think of intentions in the simple view as conforming to belief–desire pairs⁵³.

⁵² For further analysis of these matters see McCullagh (1976).

⁵³ Bratman (1984) suggests this criterion is concerned with ‘closeness of fit’ between what is represented in an intention and the action result.

The second condition amounts to the claim that the intention must be causally potent in bringing about that the world matches your intention. The condition is included to separate events that can be understood as being brought about by an individual's intention and those that just happen to them without the intention playing a role. The problem is captured in the following example. Being a fan of NYPD Blue, you have formed an intention to see the New York Police Department tomorrow. While visiting the Empire State Building, you get into an argument with an official, the result of which gets you arrested and interned overnight. In this case you did see the NYPD the next day, but we should not consider this a matter of intentional action because your intention did not play a causal role in you seeing the NYPD.

There is a familiar worry about this second condition: *aberrant causal chains*. Davidson (1963) provides the classic example. Two climbers are scaling the north face of the Eiger, attached by a rope. The first climber forms an intention to get rid of the rope attached to the second man; the forming of the intention so unnerves the first climber that he drops the rope. Now it seems this obeys both of the above conditions of the simple view, but most of us do not consider this an intentional action. The difficulty is that, even if forming the intention is causally potent in causing the dropping of the rope, it does not seem to have brought about that dropping in the right kind of way⁵⁴.

Causal accounts of intentional actions try to accommodate this worry by spelling out the 'right way' of the intention being causally connected to the action result. Many formulations have been suggested (see, for example, Mele, 1994 or Searle, 1979), and whether we agree with one or another will involve our intuitions about what counterexamples are problematic. For Davidson (1963), what is required is a direct causal relationship between the mental intention state and the real world result. We must see the pairing as non-accidental or not involving chance, or at the least, if chance is involved it must not be the only consideration as to why the event occurred as it did. This leaves open exactly what it means for an action not to involve chance or luck. In section 3.6 we will return in more detail to these considerations.

In summarising the simple view, for an action to be intentional, we must have an intention to *x*. Further, our intention to *x*, the belief–desire pair, must be causally active in bringing about *x* in the right way. The question we need to answer, then, is whether mental events will qualify as intentional actions under these conditions. This will be a useful beginning in understanding the agentic nature of mental events, for we will begin to see the difficulties of describing mental events as intentional action. One difficulty we will uncover is the access puzzle. This will be particularly problematic for understanding IBPS as an intentional action. I will begin with Galen Strawson's (2003) sceptical account of mental

⁵⁴ 7: This problem of aberrant causal chains is problematic for many accounts of intentional action; we will return to it in sections 3.2.4.

action, which can be understood as a version of the simple view, before turning to a more general interpretation of the simple view⁵⁵.

3.2.1 Strawson on Mental Action

In 'Mental Ballistics' Strawson (2003) is interested in what types of mental activity can be seen as actions. Strawson understands actions as something done by an individual, done intentionally, and done not by doing anything else. He considers most mental events as 'ballistic' in nature, i.e. not under the control of an agent, not something they do intentionally, but something that occurs involuntarily through the operation of the mind. In Strawson's view mental activity that can be described as action is limited to prompting the mind. I can intend to 'think' about cats but what comes about once I have prompted my mind to think about cats is unintentional. This prompting of the mind is a mental action, something I do, do intentionally and don't do by doing anything else. It is, then, possible for us to prompt our own thinking, to cajole the mind into bringing about new content; but importantly for Strawson, what happens after the prompting is not under our control, and so not part of an intentional action. Similarly, according to Strawson (2003) the process of the recalling of memories (a process we will investigate later), like many types of mental process, can be partially a matter of action. You can prompt the mind to recall a name, and this is an action. But mental actions go no further than this. What comes to mind, the new content, is not part of the action, because it is an involuntary activity of the mind that is not under the control of the agent. Any new content that comes to mind as a result of prompting the mind is independent of any effort or will in trying to bring that content about. The mental activity that results from prompting actions should not be seen as having anything to do with the intentional prompting action (apart from it causally resulting from that action in a non-controlled way), and particularly nothing to do with free action, because it is not under the voluntary control of the agent.

Here Strawson is making a strong claim. We might want to think, or have the intuition, that we can voluntarily bring about content. That when I try to recall the name of my 4th grade English teacher and I do manage to recall that name I intentionally bring about that content. But Strawson is suggesting that the nature of thought rules out most kinds of mental activity as being under our control in this way. Mental action is limited to a small number of action types that don't involve bringing about content. The main type of action associated with our thoughts is the catalytic prompting of setting further processes in motion. Strawson allows that there may be many types of prompting actions — imagery⁵⁶, language prompting, and so on — but all are characteristically acts of prompting. That is, they are (crucially, for Strawson) separate from any further mental activity that results from, or is caused by them.

⁵⁵ In accessing the usefulness of understanding mental events as intentional actions, I am not committed to the truth or falsity of any particular view, but only to the general causal claim that actions are caused by intentions.

⁵⁶ Strawson does not think that the bringing about of images in the mind is an intentional action but he suggests that the manipulation of an image in the mind might be an action (see footnote 19).

Another type of mental activity that is considered by Strawson to be a case of action is focusing. Focusing has to do with concentrating, of bringing the mind to bear on some relevant content. Through the action of focusing we may bring our mind back to the problem at hand, and prevent it from wandering onto other topics. Other types of mental behaviour mentioned by Strawson, that might be actions, include the holding open of a mental space in our considerations, if, for example, we think we have left out some important information, or the activity of unpacking ideas that have arrived very quickly. We can place these types of concentrative activities into a second class, which I call shepherding actions. Mental shepherding and promptings are actions because they have a mental subject with which to form an intention about. In these cases, the content (to utilise Strawson's example) 'What do I think about democracy?' is available to be shepherded or used as a prompt by an intentional agent. It is therefore possible to intend to do something with that content. For example I can prompt my mind to think about democracy or to remember a speech by Churchill or to consider the relation of democracy and access. The intentional agent can then be described as acting to prompt the mind: he is intending that something be done with that content. If he succeeds in prompting the mind, that prompting action is complete; the further activity that comes after the intentional prompt is not part of the action of prompting. Both classes of mental action, shepherding and initial prompting, share the characteristic that neither of them involve the bringing about of new content (Strawson, 2003).

Under Strawson's (2003) view, IBPS and memory recall may be a matter of action in so far as they utilise prompting. In prompting the mind to remember or examine content related to a problem, we might engage in mental prompting actions. Further, in refocusing the mind on a problem, we may be performing a mental action. Under Strawson's view, many of the activities of preparation may be a matter of prompting or shepherding action, but what happens downstream, the further stages of IBPS — even if they are causally connected — will not be part of that action, for such bringing about of content will be not be under the agent's control, not something he aims at doing intentionally⁵⁷.

We can begin to see a problem with Strawson's (2003) account by utilising an analogy with physical action. In their paper Mele and Mosler (1994) utilise the following type of example. Davide wants to play his favourite computer game, Sensible Soccer (Sensi). He knows he can switch the game on by pressing the load button on the computer and Sensi will appear on his TV screen. When Davide presses the button and Sensi appears on his TV screen he can be said to perform the action of pressing the load button, but he can also be said to perform the action of switching on Sensi, and this is the case even if part of the operation of bringing it about that Sensi appears on the screen is not under his control, i.e. it is down to the mechanical and electrical operation of the computer and the TV. This notwithstanding, Davide turning on Sensi is an action on his part.

⁵⁷ In chapter 4 we consider whether we can be morally responsible for the consequences of this type of mental action.

This type of example seems to be a dilemma for Strawson (2003). He has to concede either (1) that in the computer case, since part of this event is outside of the agent's control, it is not an action, or alternatively (2) that part of an event not being under the control of the agent will not disqualify that event from being an action, in virtue of the whole process not being controlled. It seems counter-intuitive to accept the former: the computer case does seem to be an action; but if Strawson does not accept this, then it seems his case for eliminating many types of mental events as actions, based on a lack of control, won't go through, for we have physical cases of action, (such as the computer case), within which the whole process is not controlled, and so we would require further reasons for thinking that mental action can't match this model. It seems then that we have good reason to doubt Strawson's (2003) claim that actions cannot contain involuntary ballistic elements.

Strawson's (2003) account can be seen as a version of the simple view. The computer case we explored above gave us good reason for not adopting Strawson's (2003) view. Strawson's (2003) view represents a sceptical view about mental agency, and about the extent of action in our mental lives. Under this sceptical view, IBPS would not be a matter of intentional action outside of small prompting events. If we want a comprehensive exploration of whether IBPS is an intentional action we will need to look at more liberal views of mental agency. To begin with, we can consider whether IBPS can be considered an action under a more general and less strict conception of the simple view of intentional actions.

3.2.2 Is IBPS an Intentional Action Under the Simple View?

In considering this question it will become apparent that IBPS, under the conditions of the simple view (and perhaps any view of intentional action) is not a straightforward case of an intentional action. The first thing we require, for IBPS to be an intentional action under the simple view, is that an intention state be present within the process. If there is no intention state, there is little point in us worrying whether it is present in the right way. (We will get to that consideration in a moment.) We can begin by asking whether a belief–desire pair can be ascribed to IBPS. In pursuing a particular problem within the first stage of IBPS, preparation, you have, or at least can have, pro-attitudes towards that problem. For example, you can have a desire to find an answer to your problem, or to keep on pursuing the problem, or a number of other pro-attitudes (section 3.1). To make this an intention you will also require a belief, that what you are going to do is at least one way to get what you desire.

One way an intention, based on a pro-attitude–belief pair, could be achieved in IBPS is an explicitly conscious (reportable) forming of an intention within preparation to solve your problem. Within preparation, such an intention, we should presume, is unsuccessful if the further stages of IBPS are to take place. For in preparation your rational deliberation ends without resulting in a positive answer to the question at hand. You may, if you are aware of the ins and outs of the IBPS process, realise that what is most advantageous at this time is to take a break from actively pursuing the problem, and take up the pursuit at a later date, perhaps focusing on another aspect of the problem that seemed promising. In

thinking this way, it might be possible to form an explicit intention to stop thinking about the problem now and continue thinking about the problem at a later date. This will require an understanding of the process of IBPS. The following problem highlights this need.

To be able to form an intention I need to have confidence that the intention I wish to form will bring about the action result. For example, Al wants to swallow whole a regular size cantaloupe. He has a desire to do so, but this will not be enough. Along with this desire he requires a belief that he is capable of swallowing a cantaloupe whole. In normal circumstances this will be impossible. Mele (1992c) suggests that this limitation on forming intentions can be understood in terms of a belief constraint. An individual will not be able to form an intention to x when they believe that they won't, or are unable to carry out x . The difficulty in IBPS is that the individual might believe, quite intuitively, that stopping thinking about a problem will not help him solve the problem. Indeed stopping thinking about the problem may look to the layman very unlikely to lead to future problem solving success. In order to intend to solve a problem by stopping thinking about it, I must at the very least not believe that this is impossible, and it seems intuitive that I will indeed need to have some positive confidence that the action is possible. I will therefore need confidence in the process of IBPS that will be achievable through an understanding of the process of IBPS or at the very least some first hand experience of the process will be necessary. What an individual will require is sufficient background information that the behaviour of not thinking about a problem will produce results.

A further concern may arise if intending to problem-solve in IBPS entails us having to be able to stop thinking about the problem. This last worry is similar to problems of intentionally trying to faint or go to sleep (Mele, 2009). These events are impossible to directly intentionally bring about. I cannot intend directly to faint, but (in some circumstances) I can intend to put the chloroform rag to my face. But unlike these cases, a better understanding of how an individual forms an intention within the process of IBPS will be sufficient to overcome this difficulty. It is not, in this instance, a case of having to act indirectly to bring about the situation where the individual stops thinking about the problem for the moment. Rather we could consider that the individual would form the intention to think about the problem later and the purpose of thinking about it later is to provide time for the process of incubation to produce a solution. The explicit intention will be formed from a desire to solve your problem and a belief that discontinuing thinking about it now, but continuing to think about it later, will be the best way to solve your problem. To be in a position, then, to form an intention of the sort required by the simple view, I will need a degree of confidence that my plan to stop thinking about the problem is likely to be productive and this may require some knowledge of the process of IBPS either in the form of explicit knowledge or of some working know how. This gives us a basis for the requirements that will be required for IBPS to be considered an intentional action under the simple view and will somewhat restrict the cases where IBPS can be seen as an intentional action.

3.2.3 The Access Puzzle

But there is still a deeper worry here. Will this kind of coarse-grained intention, an intention to problem solve, be sufficient to represent the action result? In the simple view, the action result has to fulfil the condition of tight fit. The end result of the action has to match the content of the intention. In IBPS, the action result is the answer to the problem being considered. But how can I form an intention to bring about this particular action result, i.e. one that represents the answer to the problem being sought, without first having access to the content of the problem solution? This problem has been called the access puzzle (Proust, 2001).

Proust (2001) introduces the puzzle in terms of the problem of attending to perceptual content. She asks how it is possible to attend to something that you have no awareness of, that is, before you have isolated that particular piece of content. Proust here is concerned with understanding attending to mental content as a mental action. For Proust not all attending to mental content is an action. Sometimes, for example when hearing a shout you attend to stimulus without acting. On other occasions however we need to try to attend to content. For example if you are at a cocktail party where there is a babble of noise and you are trying to attend to what a particular individual is saying. For Proust this is a mental action because it passes the 'try test'; it is something you can try to do. The problem with understanding this type of attending as a mental action is associated with the mental representations that are required to act. It seems that (on the simple view at least) in order to act to isolate particular content I have first to represent that content.⁵⁸

The access puzzle is a puzzle about acting to gain content or a particular mental state. It has a broad remit. Proust (2001) thinks it can be applied to beliefs, desires and all types of mental state that might be brought about by action. It is only applicable to cases of action and then only to instances where it seems the actor will require to have access to the outcome of his action for the accomplishing of that outcome to be an action. In ordering our minds we often want to get hold of content, such as recalling a memory, but this seems impossible to accomplish as a matter of intentional action because it seems I would need to have access to the content I want to retrieve, if I am to represent that content in my intention and I need to represent the content in my intention for the retrieving of content to be an intentional action.. The access puzzle, then, is a problem if we want to understand the activity of the mind in terms of intentional mental action.

In this chapter the access puzzle will be of particular interest in our investigation into whether the processes of IBPS and memory recall can be seen as intentional actions. Intentional actions that bring about content are particularly vulnerable to the access puzzle because to perform intentional actions agents are often (always on the simple view) required to represent what they want to bring about. The question

⁵⁸ We might consider this a strange way of understanding content and the process of attending. I will rephrase the access puzzle in terms of mental actions in general, where such consideration will not be in play.

that arises for any account of intentional action that demands the bringing about of new content is how such content is to be represented within the intention. It is this particular problem of representing content you wish to bring about intentionally within intentional action that I will refer to as the access puzzle.

Here the puzzle is circular:

(P1) is the assumption that you do not have X in your possession, where X is a piece of mental content.

(P2) is the claim (condition of the simple view) that you can't intentionally act to get X without already representing X in your intention to act.

(P3) is the claim that it is not possible to represent X without having X in your possession.

Conclusion: You cannot act intentionally to gain X.

The puzzle cannot be avoided by suggesting that you already possess X. In this case it seems that there is no need to act. For example if I already know the date of John Lennon's birthday, I will not need to perform an act of retrieving that content (nor may such an act be possible).

P3 will come under the most scrutiny: is it the case that you cannot represent mental content unless you already possess the content? In the case of the simple view, because of the condition of tight fit - that you need to represent x in your intention for x to be seen as an intentional action - there is no solution to the access puzzle.

3.2.4 The Consequences of the Access Puzzle for Understanding IBPS as a Simple Intentional Action

It is fundamental to the nature of IBPS that the goal of the process is to provide an answer to the question at hand, and it is a requirement of the simple view that our intention refers specifically to the action result. In order to intend to find an answer we must form an intention in preparation. But within preparation we do not know what the content of the answer will be, so we cannot have a representation of it as part of our intention, and so because of this characteristic of the simple view any answer that does result will not be intentional, because we are unable to refer to it in forming our intention in preparation⁵⁹.

It might still be said that I can intend to problem solve by having the intention of 'stopping thinking about the problem now'. But this will not be an acceptable answer for the simple view, because the condition of tight fit states that in order for an action to be intentional we require a specific intention that matches the action result. We must specifically intend to x ; in intending to solve your problem by stopping thinking about it now, in hope that this will bring about you solving the problem later, you will at best make 'thinking about the problem later' an intentional action under the condition of tight fit. What if we include the more complicated intention to problem solve by stopping thinking about the problem now

⁵⁹ Memory recall also encounters much of the same difficulties as IBPS; we will discuss memory recall in more detail when we consider more advanced causal views.

but continuing to think about it later? Since this action result can be represented in the intention at least to the extent where it can be said that I solved the problem intentionally we can avoid the access puzzle. But we may still think that no specific answer to a problem can be brought about intentionally only solving the problem can be seen as intentional. Further, the second condition of the simple view, that the intention be causally connected in the right way, may be problematic. Does forming an intention to problem solve later actually bring about you problem solving later? I will examine this worry now and return to the debate about whether a coarse-grained intention may be sufficient to describe IBPS as intentional when we look at more advanced causal accounts of intentional action (section 3.3).

3.2.5 Causal Connection within IBPS

Even allowing for explicit intention-forming within IBPS, we might still be concerned that the process of IBPS does not have the right kind of causal connections to be considered an intentional action. Within the basic four-stage model of IBPS we looked at above (sections 2.4, 3.0), the causal connections between the stages of the process seem unclear. Incubation takes over when deliberation within preparation fails to provide a product, but how is this failure causally related to the processes within incubation? In order to understand IBPS as an intentional action, preparation must be causally connected to incubation and incubation must be causally connected to revelation. We have argued in chapter 2 that the incubation effect has a positive outcome for producing revelation. This is strong evidence (as strong as the evidence for the incubation effect itself) that whatever goes on in incubation has a causal impact on revelation. We may not need to know the details of what goes on. If incubation has a measurable positive effect on revelation, then this may be enough to say that in many cases incubation causes revelation. The argument that preparation causes incubation is not so clearly supported by experimentation. If we look at the two mechanisms proposed for IBPS, both suggest that the work of preparation will be essential for the processing within incubation. In having evidence that preparation helps incubation and incubation increases the chances of revelation, it can be suggested that, at least in some instances, preparation causes incubation and incubation causes revelation (no conceivable evidence here would give us direct evidence of causation). We have then sufficient reason to think the stages are causally connected.

But this does not directly give us any evidence that an intention to problem solve will be causally effective in bringing about incubation or revelation, for as we have discussed in IBPS, the most natural case will proceed without a explicit intention being formed at the end of preparation. Even if we think that an intention ‘to problem solve’ can be formed we may wonder if it is causally connected to the bringing about of the solution. I can perhaps make this conscious intention, but as to its effectiveness I am left in the dark⁶⁰. Even if my intention to stop thinking about my problem is effective, and my stopping thinking

⁶⁰ Some of the causal difficulties discussed here may also be a difficulty for the process of imagining. Is it possible to imagine in a way that is a simple intentional action? Visual imagery is an area of mental activity that has been traditionally associated with action (Colin McGinn, 2004; Wittgenstein 1981). Consider the case of imagining of simple images. Can I intentionally conjure a specific image in my mind.

about the problem leads to me solving the problem, as we have discussed above, the requirement for a close fit will rule out the bringing about of the solution being intentional.

3.2.6 Conclusions: Mental Events and the Simple View

In conclusion, describing IBPS or memory recall as an intentional action under the simple view faces many difficulties. Firstly, in order for IBPS to be intentional we must form an intention in preparation. This is already an unnatural picture of IBPS, for we might consider that much problem solving goes on without an explicit intention being formed. Secondly, any intention that is formed will have to be coarse-grained because of the access puzzle and the fact that you do not have access to the specific outcome upfront in the content of your intention. Thirdly, it is not straightforwardly the case that such an intention will be causally connected to further problem solving stages. On Strawson's (2003) account, the upfront activity involved in memory recall and IBPS can be seen as including intentional actions of prompting or shepherding, but the bringing about of new content cannot form part of those actions. The simple view cannot accommodate an account of IBPS as an intentional action, either because of the restriction of the type of intention that can be seen to be causally active in the right way and the condition of tight fit.

The simple view can only accommodate a small range of mental events as actions, but perhaps this is just so much the worse for the simple view. For the simple view has come under a lot of criticism, particularly because it seems unable to accommodate much of the physical activity we commonly think of as actions. Take the example of walking to work. In walking to work you take many steps. Under the simple view, you need to have a specific intention to make each step for all or any one of those individual steps to be seen as intentional actions. This seems very counter-intuitive. To accommodate this problem philosophers (Searle, 1979; Bratman, 1984; and Mele and Moser, 1994) have turned to more advanced causal accounts. In so doing they have increased the range of physical activity that can be seen as intentional action. It is worth then exploring if advanced causal accounts will increase the range of mental

Take the example of a pink elephant. It might seem straightforward that I can form an intention to imagine a pink elephant and it seems a pink elephant comes to mind. Here the access puzzle is not a difficulty. I have a representation of pink elephant in my intention; we can suggest a linguistic representation and what follows is an image that matches that linguistic representation. What is not clear in this case is whether my intending brings about the image of the pink elephant. We are faced again with Strawson's objection that the mental event, in this case, bringing about the mental image, is involuntary. This becomes clearer if, as Strawson (2003) suggests, we try not to imagine a pink elephant. It seems that the image just comes to mind, at least in a lot of cases. This gives weight to his claim that the process of visual imagination is automatic in a way that prohibits intentional control. Under the simple view, you are supposed to form an intention and then act upon it. But in the case of imagination, such as imagining a pink elephant, the process of imagining takes place alongside the forming of the intention. This highlights the difficulty in seeing the bringing about of the image as causally connected in the right way to the formation of the intention. If we rule out these cases, we may still think that cases of controlling mental images are still a matter of intentional mental action. That is, if we have adequate control, we can, for example, rotate the image of the pink elephant 90 degrees. The example of conjuring mental images seems to be a case where lack of control of the causal process rules these conjuring out as intentional actions.

activity we can consider intentional action, maybe even to the extent that it will accommodate IBPS. I will turn now to these more advanced causal accounts of intentional action⁶¹.

3.3 Advanced Causal Views

Advanced causal views have been utilised by Mele (1997) to describe some types of mental events as actions. In this context, whether memory recall can be understood as a mental action has been discussed widely (see Mele, 1997, 2009; Hieronymi, 2009; Strawson, 2003). Within this section I will set out and discuss this debate, because it has interesting repercussions for whether IBPS can be understood as an intentional action.

3.3.1 Simple Problems? Searle's Answer: Two Types of Intentions

As we have seen, the simple view comes under criticism both from Davidson's problem of deviant causal chains and for being too restrictive in the types of physical activity that qualify as actions. A classic response to the problem of deviance is provided by Searle (1979). The solution also helps us overcome the restrictive nature of the simple view in defining too tightly the range of activities that can be considered intentional actions. Searle (1979) suggests the problem of deviance arises because there is a time gap between the formation of the intention and the subsequent action. This gap in timing allows for causal chains to go awry. His answer is to insist on two distinct types of intention. The first are *prior intentions* and the second *intentions in action*. Prior intentions are standard intentions as we have understood them in the simple view, intentions formed prior to the action that represent the goal state. In contrast to prior intentions, intentions in action are maintained throughout the bodily action (here Searle is speaking specifically of bodily action) and help guide the process. (We will discuss the nature of intentions in action below.) Searle (1979) suggests that for bodily motion to be intentional it must be caused by an intention in action, and this must be the proximal cause of the behaviour. In Searle's account, for an action to be intentional it need not have a prior intention, but it must have an intention in action. This allows room to include spontaneous behaviours such as answering the telephone, or walking around the office while working, to be intentional actions even although you have no prior intention to perform them. It also provides an explanation for why the Davidson's climber's dropping of the rope should not be considered an action. In this case, although the climber had a prior intention to drop the rope he has no intention in action guiding his dropping of the rope. Therefore, for Searle the behaviour should not be seen as an intentional action. We might think that because this kind of causal account does not rely on prior intentions, and it was the need for a prior intention that matched the action result that ruled out IBPS and some other mental events (memory recall) as an intentional actions under the simple view, it is worth examining whether this causal view can accommodate other mental events as intentional action. In order to do so we need to examine more closely the nature of intentions in action.

⁶¹ In the next chapter (section, 4.2) I will consider Bratman's (1984) account of action that is motivated partially by the same deficiency in the simple view.

3.3.2 The Nature of Intentions in Action

As we noted above, intentions in action are states that persist throughout the action and guide bodily movements. How is this supposed to work? A quick example will get us started. You decide to walk to work, you form an intention to walk to work, and you intentionally walk to work. The taking of a single step during the process of walking to work is an intentional action, but it is not the case that you have to form or think of a specific intention to take each step. The intention in action to walk to work occupies the causal maintaining role without the walker having to make specific distinct intentions for each separate step. It follows that your taking of an individual step is an intentional action despite the fact that you had no specific intention for the taking of that step. It seems, then, that we no longer require a one-to-one match between intention and any particular part of a process for that part to be an intentional action, as was required by the condition of tight fit in the simple view. All that we require is an intention state that is fulfilling the causal role of maintaining that process. So many different events, taking a multitude of steps on your way to work, can be understood as intentional actions without there being independent intentions to perform each step. They can each be seen as intentional because of the causally potent intention in action that continues throughout the process.

What remains to be answered about this picture is how exactly an intention in action guides the course of an action. Here Searle gives a role to the experience of acting. When I raise my arm I have an experience of raising my arm. Such an experience in action is, for Searle, a representation of that action. So we have two components: the actual physical activity and the experience of acting in that way. Both are crucial for Searle's account. This experiential representation has a role to play in providing feedback that guides the actual physical activity of the action. For Searle, for an action to be intentional, the representation in the form of your experience of acting has to be matched by the physical activity of you actually moving your arm. His account is specifically tied to physical activity, and so will not be directly applicable to mental events. Nor will an experience of mental activity necessarily be present. Can an account along these lines accommodate mental action? Mele (1997, 2009) for one, believes that it can.

3.3.3 Mele and Moser: The Advanced Causal Account

In their 1994 paper, Mele and Moser aimed to provide a complete causal account of intentional action. Their view is a specific way of developing Searle's view of intentions in action. I will call this the advanced causal account (ACA). ACA contains a range of conditions that a behavioural episode has to meet for it to count as an intentional action. Firstly, an individual's intention to perform a specific act must contain an *intention-embedded plan* to perform that act. Secondly, the intention must then follow that plan in being suitably guided by it. Finally, there are a number of conditions that attempt to exclude 'lucky events' as actions. One concern is how skilled action can be separated from lucky action. How can we distinguish between an archer hitting a bull's eye luckily or through her skill in firing arrows? For the

moment we will leave concerns of luck and skill to one side (we will pick them up again in subsection 3.4.1) and concentrate on the first two conditions noted above. As we shall see, understanding intentional action under this type of causal view will accommodate an increased range of mental activity, but in the case of IBPS the access puzzle will continue to be a problem.

3.3.4 How Intention-Embedded Plans Guide Intentional Action

What are intention-embedded plans and how do they represent the goal of the intentional action? Mele and Moser (1994) suggest that on a common notion of intentional action a person can only perform an action if she or he has some representation of that action. In our above discussion of Davidson, the representation of the intended result is contained within the belief–desire pair. I desire to drink x and believe that to drink x I must act accordingly. Here x is something in the world, like a cup of tea, and my intention contains a representation of that object, and of drinking, and perhaps of many other conditions of that drinking. Mele and Moser (1994) suggest that within intention-embedded plans an act, for example moving your arm, can be represented by a plan for moving your arm, and this may be nothing more than a general notion, e.g. when I move my arm ‘I go like this’. In more complicated examples it may be that many more parts of the plan are required to be represented by the agent, and this may require beliefs about the various means and ends involved. But it will not necessarily involve a belief that you will always be able to carry out the plan. This suggests that the following type of example can be included as an intentional action. John is trying to defuse a bomb. He only has one minute to do so and it usually takes him five. Nevertheless, he has a plan to defuse the bomb, and does so. This might be an intentional action even if he never believed he would manage to defuse the bomb. So intention-embedded plans require one to represent the goal of an action and the path through which that goal will be achieved at least in some minimal way.

Mele and Moser (1994) also require that for an action to be intentional, the intention plan must be followed. For an individual to follow a plan, they must be guided by that plan. It is not enough that their behaviour conforms to the plan. They draw the distinction between complying with a plan and following it. This is similar to the case of complying with a rule and following a rule. A rule can be complied with by an individual merely by chance: what is required to follow a rule is that an individual’s behaviour is influenced by that rule. You may play a game of football without touching the ball with your hands when it is in play. In so doing you abide by, or comply with, the rule. But this does not entail that you followed the rule. In order to follow a rule you must know what the rule is and this must influence your behaviour.

What does guidance mean in terms of intentions with embedded action plans? Mele and Moser (1994) state that in the case of physical action, guidance depends upon monitoring progress towards the final goal. In the case of a golfer making a stroke, his intention to hit the ball carries on throughout his bodily movement, e.g. moving the club back, reaching the end of his backswing, hitting the ball, following through. If at any time he wanted to stop his swing he could, for instance if he changes his mind about which club to use. He continues to act through the swing because he intends to do so, and this activity is

guided by his mental plan for so doing. In their view, this will require that the action ends when the guiding intention ends. Or to put it another way, the guiding role of intentions carries on until the end of the action⁶². Further, Mele and Moser (1994) suggest that a necessary condition of an intention-embedded plan playing a guiding role is that the intention, within which the plan is embedded, plays a causal role in the production of the action. This is a basic condition of any causal account. Further still, they suggest that for an action to be intentional the plan must be followed suitably closely. This is a vague term, and Mele and Moser suggest that the vague nature of the term captures the vague nature of actions. They provide an example that supports this intuition. A stunt diver lands in a baseball field three feet behind home plate and performs a somersault onto that plate. In the first example the diver is assisted by a back wind that he underestimates by 0.5 mph. In the second case he misjudges the wind speed by 50 mph. Mele and Moser (1994) suggest that example one is an intentional action where the second example is not. This is the case because in the first case the plan is followed suitably closely and in the second it is not. They suggest that whether a plan is suitably followed or not will depend on intuitions and commonsense judgements, and these judgments might depend on individual agents' evidential context (Mele and Moser, 1994). So for an intention to guide a physical action it must cause the action and be suitably guided by the embedded action plan. What about mental actions? Can they be guided in the same way? We turn now to Mele's (1997) attempt to adapt this type of causal view to mental action.

3.3.5 Mele's General Causal Theory and Mental Actions

Mele, in his 1997 paper 'Agency and Mental Action', suggests that there is no particular reason why mental actions cannot be accommodated under a General Causal Theory (GCT). He suggests that an instance ACA outlined above will be instance of GCT. GCT is included to try to accommodate mental actions. If mental actions are to abide by GCT we need to ask the following questions. How are mental actions represented and what role does that representation have in producing an action result? In terms of ACA how can we have intention-embedded plans for producing them? And how are mental actions guided by those plans?

Under Mele's (1997) account of GCT, an action can be said to be intentional if it has a certain type of causal history. The nature of the causal history of an event, such as raising your arm, will determine if it is or is not an action. No physical process that produces the same result (for example, an advanced neurosurgeon causing your arm to rise by direct brain-tampering) can be said to be intentional unless it possesses the right causal history. In the case of intentional actions, what Mele (1997) calls *pertinent mental states* must be included in the causal chain in the right way. For Mele the list of items

⁶² Some action might include further involuntary elements; we will take up this point later when we address whether IBPS can be seen as an action under an advanced causal view and whether it contains ballistic elements.

that qualify as pertinent mental states include belief–desire pairs that can be thought of as reasons for action, intentions or events of intentional acquisition. According to GCT an action will be intentional if it contains a pertinent mental state as a significant cause. (By significant cause Mele indicates that the intention need not be seen as the only cause, or perhaps the most significant cause, but it must at least be seen as a partial cause of the action.) It will also be important that the intention does not just cause the action, but that its presence sustains the action in a similar way to the role suggested for intention-embedded plans. In the case of mental action, what will play the sustaining role is the agent trying to achieve his goal.

He gives the example of trying to remember a list of seven animals whose names begin with the letter ‘g’. According to GCT we should consider this an action if the event is constituted by your trying to acquire this mental list and your trying to do this is a distinct intentional episode (Mele, 1997). What does Mele mean by trying? We can get a grip on trying by asking ourselves Wittgenstein’s (1953) question about raising your arm. What is left over if I subtract the fact that my arm went up from the fact that I raised my arm? Mele suggests that one answer is my trying to raise my arm. For Mele, trying is constituted by the extent of the effort involved in whatever it takes to bring about the end result of a physical action. (This follows McCann’s (1974) point that trying is equivalent to effort involved in bringing about an action). When I sign my name, trying is involved; it extends throughout all the period of activity required to sign my name, even if I encounter no special resistance in doing so. Trying, then, goes on throughout the whole of an action. What will count as an episode of trying will be delineated by this physical/mental activity. For Mele (1997) this period of activity begins when the individual acquires his or her intention to act and that intention begins to play its functional role of maintaining and sustaining the action. The activity of the intention can be seen as a neurophysiological event that has the causal history appropriate for a pertinent mental state under GCT.

What about cases of mental trying? Are there particular examples that are not connected to bodily action? Mele (1997) suggests that what is left over when we subtract what it is for me to have a particular mental state from what it is for me to actively obtain that mental state is my trying to obtain that mental state. On this account, some mental states will be passive- thoughts just coming to mind, e.g. what you think about your neighbour’s new car - and some will be a matter of action. For Mele (1997) the difference between passive behaviour and mental actions is the trying to bring the latter about: mental trying consists of a neurophysiological event that guides and maintains the mental activity through the presence of a pertinent mental state such as an intention. It is a matter of maintaining the process through the interaction between causally potent mental states and the online mental activity that brings about content.

On this account the recalling of seven animal names beginning with ‘g’ may be a mental action, if your trying to remember the seven animal names is causally connected in the right way to you remembering seven animal names. That is, your trying has a role in maintaining the activity of

remembering; it lasts throughout the process and has a causal role in bringing about the result of the action, in this case the new content of seven animal names. On first look, this kind of account may help us get round the access puzzle, for what seems to matter for a piece of new content to be brought about intentionally is that the trying to bring it about has a role in the bringing about of the new content. This claim that this account of mental action gets round the access puzzle requires further scrutiny. We can begin with an objection to this kind of account of mental action that is raised by Galen Strawson.

3.3.6 Strawson on Mental Action

We set out Strawson's (2003) view above (subsection 3.2.1). Strawson's (2003) view, then, is in direct opposition to Mele (1997). For Strawson, the action that is involved in memory recall is only upfront: it is the catalytic action of setting the process in motion. Any content that arises because of this prompting is involuntary activity that should not be seen as an effortful activity on the part of the agent. Trying, in this view, would be limited to the effort of prompting. Under Strawson's (2003) view, IBPS and memory recall may contain prompting actions. In prompting the mind to remember or to examine content related to a problem we might engage in mental actions. Further, in refocusing the mind on a problem we may be performing a mental action. On this view, many of the activities of preparation may be a matter of a prompting or shepherding action but what happens downstream - the further stages of IBPS - will not be a matter of action even if they are causally connected. For such bringing about of content will be purely ballistic or involuntary and not under the agent's control. The same will be true for the recalling of content. As we noted above, there are different intuitions about whether ballistic involuntary activity can be included within actions. In the computer game case (section 3.2.1) our intuitions pointed us to accept it as a physical action; here intuitions about what is or is not an action are playing a central role in our argument. Even if we side with Mele and Moser (1994) in thinking that the computer case is an action, we may still have reason to think that with memory recall and IBPS there is an additional problem in considering them intentional actions. In the computer case Davide has a representation of what he wants to bring about, i.e. the computer game Sensi is specifically represented in his intention-embedded plan (as Mele and Moser (1994) would have for physical action, and as he suggested for mental action in his 1997 paper; see section 3.3.5 above). But in the case of memory recall and IBPS no such specific representation of the content is in place. The problem here is the access puzzle, and in his 2009 paper 'Mental Action: A Case Study' Mele makes a concession based on this problem. He suggests that although the bringing about of remembering content is an intentional action under GCT recalling specific content is not.

3.3.7 Memory Recall as an Intentional Action

Mele's (1997 and 2009) primary example of intentional mental action under GCT is memory recall. His claim that GCT can provide an account that is useful for understanding mental action rests partially on the success of understanding memory recall as an intentional action. I will suggest that the application for

GCT to memory recall is restricted in a way that is important for understanding why IBPS will not be an intentional action under GCT. This requires us to look in detail at Mele's (2009) account of memory recall as an example of an intentional action.

Before we proceed to analyse Mele's (2009) account of memory recall as an intentional action we need to say a few things about the process. Within the science of memory a distinction can be made between long-term and short-term memory. Long-term memory initially utilises the same process as short-term memory in the laying down of memories and in retrieval, but a further distinct process then takes place (Bailey et al., 1996). In advocating memory recall as an example of mental intentional action, it is long-term recall that is my primary target. Remembering seven animal names beginning with 'g' is an example of the type of recall I wish to discuss. Here is the process described at the personal level by James Mill (1869). There is, he says,

a state of mind familiar to all men, in which we are said to remember. In this state it is certain we have not in the mind the idea which we are trying to have in it. How is it, then, that we proceed in the course of our endeavour, to procure its introduction into the mind? If we have not the idea itself, we have certain ideas connected with it. We run over those ideas, one after another, in hopes that some one of them will suggest the idea we are in quest of; and if any one of them does, it is always one so connected with it as to call it up in the way of association. I meet an old acquaintance, whose name I do not remember, and wish to recollect. I run over a number of names, in hopes that some of them may be associated with the idea of the individual. I think of all the circumstances in which I have seen him engaged; the time when I knew him, the persons along with whom I knew him, the things he did, or the things he suffered; and, if I chance upon any idea with which the name is associated, then immediately I have the recollection; if not, my pursuit of it is vain. (Mill 1869)

Here is what William James has to say about it:

In short, we make search in our memory for a forgotten idea, just as we rummage our house for a lost object. In both cases we visit what seems to us the probable *neighbourhood* of that which we miss. We turn over the things under which, or within which, or alongside of which, it may possibly be; and if it lies near them, it soon comes to view. (James 1890)

In both of these cases we have memory recall described utilising the language of action. But as Strawson (2003) notes, both active and passive language can be used to describe cases of mental activity: because most events can be described in both ways the language here will be unhelpful in arriving at the appropriate nature of the mental activity in question. Further, in the context of causal theories of action, language does not play a direct role in the attribution of action. The above descriptions do, however, help us focus on the exact phenomena we are investigating. Let us return to our remembering seven animal names beginning with the letter 'g'. I will refer to the activity of an individual remembering seven animals beginning with 'g', as Mele (2009) does as 7-s (they 7-s). Mele (2009) makes the following four claims:

- (1) Your thinking of the name 'goat' or any specific content is not an intentional action.

- (2) Your thinking of seven animal names starting with ‘g’ is not an intentional action.
- (3) Your trying to bring it about that you 7-s is an intentional action.
- (4) Your bringing it about that you 7-s is an intentional action.

Let us take (3) first because it is the easiest. Here the claim is that your trying to bring it about that you 7-s is an intentional action. This can be seen in several lights, depending on what is meant by trying, by Strawson (2005) as a prompt or by Mele (1997) and McCann (1974) as the activity of going about an action. In both cases, whether it can be described as an intentional action will depend on the causal relation between how you try to bring about x and your x -ing. In Strawson’s (2003) view, it is under your intentional control in accordance with the simple view (that you must form a specific intention to perform that action and that intention will be causally necessary for bringing about the result of your actions), and for Mele (2009) that it is under your intentional control according to GCT(that an event can be said to be an intentional action if it has a pertinent mental item as a significant cause). Under both accounts, (3) - the claim that your trying to bring it about that you 7-s is an intentional action- — will be correct.

Let us move on to (1), (2) and (4): since they are related I will deal with them together. In (2), your thinking of seven animal names starting with the letter ‘g’ is not an intentional action is a rejection of the claim Mele makes in his 1997 paper,

that if by so focusing your so doing results in your acquiring a conscious mental list of seven suitable names. Then you have performed an action describable as ‘thinking of seven kinds of animals having a name starting with the letter ‘g’’. (Mele 1997)

The view has been modified in light of Strawson’s (2005) objection that new content cannot be the subject of an intentional action. Mele (2009) now suggests that ‘Your thinking of seven animal names starting with ‘g’ is not an intentional action. He claims, however, that (4) is correct, ‘Your bringing it about that you 7-s is an intentional action’. What for Mele is the problem with you 7-s being an intentional action? The problem rests, as is now familiar, on whether the specific names of the seven animals you remember are represented in your intention to remember seven animal names⁶³.

The difficulty boils down to two interpretations of how we should describe what you are doing when you say you are thinking of seven animal names, whether your bringing it about that you 7-s is connected to trying events. Consider Mele’s (2009) example of trying to sneeze. It may be very difficult, if not impossible, to intentionally try to sneeze, but it is quite easy to try to bring it about that you sneeze (just

⁶³ . This is, unlike Mele’s (1997) claim and not in direct contradiction with Strawson’s (2003) main claim that any bringing about of content cannot be an intentional action and claim (1) made by Mele (2009). Your thinking of a ‘goat’ or any specific content is not an intentional action..

reach for the pepper or interact with a cat)⁶⁴. So bringing it about that you 7-s can be directly connected to trying events, but the same cannot be said for thinking of seven animal names, because thinking of seven animal names refers to events that you do not *directly* try to do, but can only bring about by doing other things. In making this distinction, Mele (2010) provides for a type of mental action, but it does not help us to deal directly with the difficulties that arise from the access puzzle. In our intention to act we will still be required to specifically represent new content if that new content is to be describable as part of an intentional action. This becomes clearer when we ask the question, 'Can a person intentionally bring it about that they recall a single animal name beginning with 'g'?' Well, why not? Remember Mele's reason for (1) 'Your thinking of a 'goat' or any specific content is not an intentional action' being unintentional and Strawson's claim that a specific content coming to mind is never an intentional action. If 'goat' just springs to mind immediately without being linked to a trying event, then for Mele (2009) it is not in any way intentional. Mele has overcome this by making a distinction based on trying, and at base trying is intentional because of it agreeing with GCT. That is trying events can be seen as episodes of distinct intentional activity because they are guided by a mental representation of how particular events should be brought about. The case of you 7-ed is intentional because you have a plan to bring about your remembering 7 animal names beginning with 'g' that consists of you doing what you normally do when you try to remember a list of names and this plan is a distinct mental episode that is casually important in you bringing it about that you 7-ed. The case, if it holds up, also holds for single cases of memory recall that can properly be described as tryings. The process of the recalling of memories can then be seen as an intentional action as long as the process involves the trying to bring about content.

But even if the process of memory recall can be understood as an intentional action, that is, the bringing about that you 7-s (or the recalling a single name 1-s) is an intentional action. The bringing about of the actual content of those seven names, e.g. gorilla, giraffe, groundhog, etc (or of a single name), cannot be described as intentional or as part of an intentional action or as an intentional action. The problem becomes clear when we realise that a wide range of new content can fill the gap. In an adaptation of Mele's (2009) example, if you try to remember an animal name beginning with 'g', any specific name that begins with 'g', e.g. goat, will match the representation present within the intention. The exact content that comes to mind then seems not to be controlled by the intention even if the bringing about that you recall an animal name beginning with 'g' is. Your reasons for bringing about the new content are neutral between the different possible contents that fulfil the representation of your intention. This being the case, we cannot say that the production of any specific piece of content is intentional, but only the trying to bring about of a product. What rules out the specifically bringing about of content as an intentional action

⁶⁴ This distinction between intentionally trying to bring something about and intentional bringing something about, is similar to Hieronymi's (2009) distinction between intentional control and managerial control, where managerial control is the ability to bring about events intentionally by indirect means (see subsection 5.4.1).

in (1) above is the access puzzle. It is the fact that the representation of the action cannot contain a reference to the specific content that actually arises from the intention that prevents any content brought about by memory recall being describable as part of an intentional action. In (2), what rules out the coming to mind of seven animal names being an action is the same, in that each of these episodes is not described in the intention; but this fact does not rule out the description (4) being intentional, because this description picks out an episode that is represented by your intention. We might consider the inability to bring about specific content intentionally trivial but if we think that this activity is not connected to our reasons for acting then we should not consider it agentive. We will return to considerations of the importance of bringing about specific content in an agentive way in the next chapter (section 4.6), but we can turn now to IBPS and see whether any of the events describable as actions within memory recall help provide an account of intentional action for IBPS.

3.4 Can IBPS be Understood as a Mental Action under GCT?

From our analysis of the access problem, it seems that any new content brought about by IBPS will not be part of an intentional action. That is, specific content cannot be brought about intentionally unless we can represent it up front within our intention. We may think that this is already a very good reason for looking for an account of mental agency for IBPS that is not based on framing IBPS as an intentional action. Before we do that, let's take the time to see if the bringing about of an answer (not a specific answer) through the process of IBPS can be seen as an intentional action under GCT. When discussing the simple view above we proposed an account of intentional action for IBPS. In that proposal an explicit intention was formed within preparation to stop thinking about the problem now but to continue thinking about the problem later. Here we had the belief, because of our understanding of the process of IBPS, that stopping thinking about the problem for the moment but thinking about it at a later date would bring about an answer to the problem in the form of a revelation.

Under GCT, what is necessary for a mental event to be an intentional action is that we can be said to try to bring about the event, and that this trying is a distinct mental episode that is linked to a specific neurological event that persists throughout the course of, and maintains, that action. IBPS seems distinct from memory recall in that there is no continuous period of concentration spanning the process as a whole that can be associated with a trying event. Here we can return to the question of whether actions that depend on trying, like holding up your arm or recalling a memory, depend on a continued exertion on your part (sections 3.3.4 and 3.3.5). Do these actions end when the exertion ends or can further future activity count as part of the action? If trying actions stop when the exertion stops, then it may seem to be the case that IBPS will not be an intentional action, for in incubation we are not concentrating or focusing on a problem, and it is exactly this not-focusing that seems to pay dividends in bringing about revelation.

Incubation, then, is a distinct period that we might have difficulty conceiving of as part of a trying event. Is the game over for understanding the process of IBPS as a whole as an intentional action under GCT?

As we have suggested when dealing with physical action, the fact that a period of an action (in the case of IBPS incubation) is beyond our control does not always bear on the case, for there are many types of physical action (such as the computer case, section 3.2.1) where we cannot exhibit control throughout the action.

The difficulty here is understanding what is necessary for describing an activity as part of a trying event. Intuitively we might reflect that it requires some conscious effort on the part of the person trying. But under Mele's description a trying is just the activity required to achieve some end that is maintained by some pertinent mental state, e.g. an intention embedded plan. So what it means to try to remember seems to be the same as the activity of the events that constitute remembering being maintained by an intention embedded plan.

If we apply this to IBPS, what it is to try to problem solve will be identical with all the activity of problem solving plus a maintaining intention state. When we look at the mental activity of IBPS there seem to be distinct stages of processing in play. We might then think that only part of the process, the part that can be associated with some form of effort (in this case preparation), should be linked with a trying event. On the other hand, since all the activity of preparation and incubation can be seen as necessary for bringing about an answer to the problem under consideration, we might consider that preparation, incubation and revelation and to a certain extent verification, if we consider that forming an answer will require some work utilising the new insight, can be seen as trying to bring about an answer to the question at hand, and so part of an overall intentional action. If preparation alone is suitably maintained, then we need to reassess whether further non-voluntary or ballistic operations can be seen as part of a mental action on Mele's account. We have, then, two ways forward. We can try to find an intention state that maintains the full process of IBPS (I'll call this possibility 1) or we can limit Mele-like trying events to part of the process of IBPS for instance preparation, and then try to provide an account of why the involuntary, ballistic remainder of the operation, incubation, revelation should be considered part of the intentional action (I'll call this possibility 2).

In considering possibility 1 the best candidate for such a guiding state would be an explicit intention. This would bring us back to something like the advance causal account (sections 3.3.3 and 3.3.4). Can an intention to problem solve be seen as including an intention-embedded plan that is not necessarily linked to an event of mental trying? To think so, we must think that in having an intention to problem solve the resolving of the problem at hand is represented in some minimal way within our intention. On ACA (Mele and Moser, 1994), what would be necessary for 'arriving at an answer' to be intentionally brought about is that the event be guided by an intention-embedded plan. In this case we do seem to have a plan for action, in that we will stop thinking about a problem, and by so doing allow the process of incubation to take over and provide an insight. For Mele, what will be further required is that

this plan be maintained by our intention, and this amounts to the requirement that the intention should be causally potent in bringing about the insight and that the embedded plan should be closely followed.

Well, if it is the case that we can stop thinking about something intentionally and that providing an episode of incubation by stopping thinking about a problem now to provide an insight later has a causal effect in bring about insights, then we should think that the intention to stop thinking about a problem is causally important in bringing about revelation. This meets Mele's first condition required for an intention to guide the process, that a causal connection is present. As we have discussed, whether a plan is suitably followed is vague. No clear criteria for when a plan is suitably followed is put forward by Mele and Moser (1994).

To assess possibility 1 will involve us asking to what extent luck can play a role in the fulfilment of intention-embedded plans. To assess possibility 2 we will have to consider whether trying actions can involve ballistic components.

3.4.1 Luck, Ballistic Parts and IBPS

There is now a wide range of difficulties with this view of intentionally bringing about the answering of the problem in IBPS. As we have already suggested, the number of cases where this is likely to be applicable are very small. For we will not frequently form conscious intentions to problem solve. Further, we may think that there is a problem that the control involved in the process is too weak. In ACA, the fact that the intention is causally active in bringing about an action result is sufficient for it to be understood as maintaining the process. It might be that our intention to problem solve, by intending to think about the problem at a later date, is causally potent, but as we have seen in the case of aberrant causal chains (section 3.2), there is a basic difficulty within action accounts of equivocating control and causation. Often causal connections bring about results that don't seem to be under our control. This raises two considerations: whether the extent to which luck can be involved in fulfilling intention-embedded plans (a difficulty associated with possibility 1), and secondly, whether intentional actions based on trying can contain ballistic, involuntary elements (a difficulty associated with possibility 2). In this first consideration, concerning possibility 1, depending on the conditions we include to eliminate luck, some episodes of IBPS might qualify as a possible intentional action under the advanced causal theory. I will not go through all the different ways that have been utilised to eliminate luck, because even if we succeed in providing an account of 'solving a problem' as an intentional action, we will not be able to provide an account that will accommodate the new content produced by the process of IBPS as an intentional action. It will suffice to show that understanding the character of IBPS may make an account of disqualifying luck particularly challenging. Along the way the presence of a ballistic element within IBPS will be discussed (a difficulty associated with possibility 2).

Let's return to the computer example (section 3.2.1) as a way of beginning to examine the influence of luck in IBPS. To examine this question we need a better understanding of luck. Duncan

Pritchard puts forward a comprehensive and influential account of luck in epistemology. Pritchard (2007) suggests that luck should be understood modally. An event is lucky if it occurs in the actual world but does not occur in most of the nearby possible worlds. (Here nearby is spelt out in the number of differences between the actual world and the possible world in question.) Further, for an event to be lucky it must be significant (either positively or negatively) for the individual in question. This way of looking at luck allows us to accommodate the issue of control. An event that an individual has control over will be likely to occur in most of the nearby possible worlds. Such an event can be described as *safe*. On the other hand, events that are not controlled (or not well controlled) will not occur in nearby possible worlds and so should be considered unsafe. We can for the moment understand causal connection being of the right kind to count as important for the ascription of intentional action if they are safe in the sense just outlined⁶⁵.

In the computer example, you intend to switch on your favourite game by pushing the on button. It can be correctly said that your making the game appear on the screen is an intentional action. In Mele and Moser's (1994) account their explanation for this is that there is an expectation of regularity between you pushing the button and the game coming on. We can describe this in terms of safety. We can say that if the computer works in that way in most possible worlds then the switching on of the game is intentional. It is intuitive to think that in the computer game case, if the game coming on occurred in less than half the possible worlds where the button was pressed, it would be difficult to consider that it was an intentional action. Compare this to taking a free throw in basketball. It seems intuitive that if I intend to score a basket and I, for example, attend to that attempt, concentrate, and so on, and score a basket, it should count as an intentional action on my part even if the event was unlikely, that is, did not happen in the majority of possible worlds. Mele and Moser (1994) suggest that in this sort of case there will come a point where the result is so unlikely that it does not count as an action, say, for example, less than 0.1 %. But for Mele and Moser (1994), the point at which this is reached cannot be determined. In this sense, in borderline cases whether something is or is not an action may not be decidable in Mele and Moser's (1994) account.

The difference in the computer case and the basketball case may be down to two contributing factors: (1) skill or control, and (2) being able to observe the causal influence of the part of the action that is under your control. In the basketball case, the way the ball leaves your hand is an important indicator of whether you will be successful in scoring a basket, and is itself an indicator of skill. Our estimation of whether an individual is acting when they score a basket may be influenced by the nature of how they approach the shot. If an individual approaches a shot in the right way (stands correctly, takes their time, concentrates, and so on), we may be more inclined to say that the shot was not lucky, even if when taking free throws the player normally follows this technique and is still woeful. In the case of the computer, little

⁶⁵ Although Pritchard's view provides a useful way of understanding luck, he is concerned primarily with attribution of luck to the formation of beliefs, and there may therefore still be significant problems in accommodating this account into accounts of intentional action.

skill or control is utilised in pressing the on button. So in the computer case, if the game came on only 50 per cent of the time, due perhaps to a loose wire, then we might not want to characterise the computer case as an action, and this intuition may be due to the fact that what it was in our power to do was very simple and how skilfully we performed the act could not influence further the causal structure of the process. The fact that we have no access to the further goings-on, and no definitive understanding of the process, increases our resistance to call the event an intentional action.

Consider a case where you receive coins for getting a pinball in a hole by utilising the starting spring launcher. You release the lever and move the ball up the ramp; if the weight is just right, the ball ends its arch and falls straight into a scoring hole in the machine and the coins fall out. Weighting the ball right is a skilful endeavour. Now consider the same case with an enclosed pinball machine, where the activity of the ball cannot be observed. Because we cannot observe the effect of your pinging the ball, and the direct effect of your skilled activity, but only the result of the coins falling, we may be less inclined to consider this an action, especially in cases where the task is difficult and you succeed only in a small percentage of attempts. But if we can see the mechanism and are able to observe, that, yes indeed, pinging the ball in a certain way is required to get the coins to fall, then we are more likely to consider this an action. Here the involvement of skill and our awareness that the utilisation of that skill is a deciding factor in an event occurring increases our confidence in the intuitions that a particular event is an action. Our scepticism about whether an event is an action increases with the increased difficulty of the task relative to an individual's skill level. The more difficult the task compared to an individual's skill level the less likely we are to consider it an action on their part. Or to say it another way, our confidence in the intuition that an event is an action increases the more likely it is that we consider an individual is capable of completing the intended action. On the other hand, our scepticism decreases the more we consider that skill is involved and our estimation of skill is increased by our ability to understand and observe the causal structure of an action.

It seems that the case of IBPS is more like the computer case than the basketball case in terms of our access to the further causal story, since the process of incubation is relatively unknown we can't be certain about any possible involvement of skill. There is, then, an important difference in the allocation of skill in IBPS and the basketball case.

On the other hand, for IBPS to match the computer case we will require a reliable link between forming an intention in preparation and the revelation stage. There would need to be a reliable link between choosing to think about a problem later and coming up with a new insight. The scientific evidence suggests that there is a significant advantage in adopting this technique, but whether it is sufficient, in the case of IBPS, to think that this increased probability should entitle us to talk about the solving of a problem utilising IBPS as an action is uncertain. This may be a matter of intuitions, and it may be, as Mele and Moser suggest, that there is just no definitive answer in these cases. Just like in the random computer case, part of the problem arises because of the lack of access to a particular downstream

stage that we want to consider as part of the action. In the case of IBPS, incubation is the problem. Even if incubation is casually effective in bringing about revelation, as indicated by the experiments discussed (section 3.2.4), without access to the goings-on of incubation, the difficulty of successfully describing the process of IBPS as an action increases. In the computer case, lack of access is alleviated by a strong correlation and an understanding that at least some experts (if not the actor himself) can fill in the causal story of how the computer works. So we can give a very complete causal story even if it is not accessible to the actor. But in the case of IBPS, neither the actor nor the expert can fill in the causal story, this is due to the limitation of first-person access to the process of incubation and due to the limited psychological and scientific understanding of IBPS⁶⁶. An increased understanding of the processes involved in memory recall as opposed to IBPS makes the intuition that instances of memory recall are actions stronger. The fact that the incubation effect is not well understood at all, and that the causal process behind it is still a mystery, has a negative impact on our intuitions about whether IBPS is an action. This provides us with a general conclusion about describing events as intentional actions. I suggest that the greater the understanding of causal processes involved the easier it will be to see an event as an action. In cases of mental action, a lack of understanding of the causal process involved, and in certain cases a lack of first-person access to the process, will be detrimental in any attempt to describe those processes as actions. Further, since it is also hard to access the role skill may play in incubation, the scepticism that arise from lack of access will not be placated, as would be the case if skills were in play, but increased.

In regard to possibility 1 and an account of IBPS as an intentional action that is maintained by a guiding intention state, the stage of incubation makes it difficult to assess whether the action is appropriately maintained by an intention state or whether luck plays a major role. In regard to possibility 2 if we try understand IBPS as an intentional action that includes incubation and revelation as ballistic elements, then the gap in our understanding of these stages makes any consideration of whether these stages are reliable enough to consider the whole process an intentional action difficult to assess. In such situations no positive ascription of the process of IBPS being an intentional action is appropriate.

3.5 Conclusions

In pursuing an account of intentional action for mental events, our goal has been to capture an appropriate notion of the agency involved in mental activity. The activities we have concerned ourselves with, in the main, are memory recall and IBPS. In trying to accommodate these activities into accounts of intentional action, we have progressed from the confines of the simple view, including Strawson's hardline account, through the advanced causal accounts and GCT. On no account will the bringing about of specific content in IBPS or memory recall be an intentional action. Coarse-grained accounts of intentional action for the process of IBPS, for example, intending to find an answer to a problem (subsection 3.2.4), also face many

⁶⁶ If our understanding improves then it may be that IBPS can be seen as an intentional action, but it will still not be clear that an account of intuitional action for the specific content will be possible.

difficulties in being described as intentional actions both under the simple view (subsections, 3.2.2, 3.2.3 and 3.3.4) and under GCT (section 3.4). The difficulties of access and reliability (section 3.4) provide us with good reasons for holding the intuition that the process IBPS should not be understood as an intentional action either under ACA or GCT. These difficulties, added to difficulties involving control and the unnatural nature of the account (that problem solvers will not always form intentions to problem solve or have a working knowledge of the process to make such intentions effective - sections 3.2.3 and 3.2.4), provide a substantial problem for seeing the process of IBPS as an intentional action. In the next chapter I will consider if we can be held morally responsible for the products of IBPS as a consequence of parts of that process being intentional actions.

Chapter 4

Responsibility for the Consequences of Mental Actions

4.0 Introduction

In the last chapter we examined whether IBPS could be understood as an intentional action. No such account and therefore no account of agency for IBPS based on intentional action was forthcoming. Chapter 4 continues the exploration for an account of agency for IBPS by addressing the question of whether we understand stage 1 of IBPS, preparation, as an intentional action and whether we can be morally responsible for products of IBPS as a consequence of those actions? In examining this question I will examine the related question of whether, on Bratman's (1984) account of intentional action, the products of IBPS can be seen as intentional consequences of intentional actions.

The question of whether actors are responsible for the consequences of their actions is a familiar one in the field of actions and moral responsibility.⁶⁷ The question we are concerned with here is under what conditions an individual can be seen as morally responsible, as being an appropriate target of praise or blame for the consequences of their actions. An example will get us started. Barry is playing football. He tackles Ted as he intended, taking the ball first, but in doing so he causes Ted to fall to the ground, where he breaks his wrist trying to blunt the fall. Is Barry morally responsible for Ted's wrist breaking? Whether he is or not will depend on the circumstances of the action of Barry tackling Ted. In this chapter I want to explore whether we can be morally responsible for the products of IBPS if we consider part of that process, preparation, an action.

4.1 Moral Responsibility and Causal Responsibility

To begin we need to get clear about what responsibility, and specifically moral responsibility, means. You come home from work to discover that the Acer tree in your garden planted by your father has fallen down. Consider the difference in your reactions to learning that the tree fell down after your neighbour poisoned the tree in spite, or that it had finally succumbed to the prevailing summer westerlies. We may well feel anger and resentment towards our neighbour, but, intuitively, it seems that we can't do the same for the westerly wind. According to many accounts of moral responsibility your neighbour is a possible candidate for being morally responsible, because he is a valid target for such attitudes as praise and blame, resentment, anger, or any other of what Peter Strawson (1962) calls the *reactive attitudes*. People, unlike trees and winds, are appropriate candidates in that they are able to hold and understand those attitudes.

⁶⁷ It has often entered into the debate concerning alternative possibilities and the requirement of such possibilities for genuine freedom (see Frankfurt (1969), Fischer and Ravizza (1998)). The notion here is not necessarily connected with freedom, but rather amounts to a compatibilist view that what is necessary for moral responsibility is control of specific causal events. I will not discuss responsibility for action consequences in this context.

Here we can make a distinction between moral responsibility and causal responsibility. There is a sense in which we can say the wind was responsible for blowing down the tree. If the wind was the cause of the tree falling then the wind can be said to be *causally responsible* for the tree falling. Your neighbour (in poisoning the tree, and the tree falling in virtue of that poisoning) would be responsible in this causal sense but he would also be responsible in a deeper sense, in that he is a subject that it would be appropriate to praise or blame. If circumstances prevail within which he is an appropriate target of those attitudes he will be not only causally responsible but also morally responsible.

This suggests rightly, that we will (most probably) not be morally responsible for all the things we are causally responsible for. Consider the bizarre case where your neighbour had been hypnotized to poison your tree. We might feel in these circumstances that your neighbour is not after all morally responsible (even although he is causally responsible), and this will depend on us no longer considering him, under these particular circumstances, being an appropriate target for the reactive attitudes. The question that arises is what makes it the case that people are sometimes appropriate targets of our reactive attitudes for the events they are causally responsible for and other times not.

Two alleviating conditions have been commonly accepted: the conditions of force and foreseeability.⁶⁸ The condition of force suggests that you can't be morally responsible for an event that is forced upon you. Intuitively, if you open the bank safe at gunpoint you are not morally responsible for doing so. No attempt here will be made to give an exact definition of when and how much or what kind of force precludes moral responsibility. Knowledge (in a very weak sense) can also play a role in precluding moral responsibility. Consider the case where a pilot, who after running through all the appropriate checks, starts the engines of his plane, causing the plane to explode (because of a fault in the engine). He will be causally responsible for the plane exploding, but, intuitively again, we may feel that it is inappropriate to hold him morally responsible for the explosion. This seems to be the case because he did not know, nor could he have reasonably been expected to foresee, that starting the plane would cause it to explode. This condition of foreseeability will be particularly important for ascribing moral responsibility for action consequences, for the consequences of our actions are often unforeseen. This example also brings out the fact that in particular cases it may be difficult to assess whether an individual is morally responsible for an outcome, since this will depend on a judgment about what is reasonable, and such judgments will depend on the context in which the event occurs.

4.1.1 Ascribing Moral Responsibility to Actions and Events

We have considered two negative conditions that exclude individuals from being appropriate targets of the reactive attitudes, and therefore of moral responsibility. Are there any positive conditions for the

⁶⁸ These conditions come originally from Aristotle (2009).

ascription of moral responsibility? One influential account is given by Fischer and Ravizza (1998). For Fischer and Ravizza, the controlled bringing about of events, by agents, is required for ascriptions of moral responsibility, where to bring about events in a controlled way individuals have *to exhibit guidance control*. For an individual to exhibit guidance control they have to be *moderately reasons-responsive*, and the mechanism that issues the event has to be their own. What it means for a mechanism to belong to or be owned by an individual is based, in Fischer and Ravizza's (1998) account, on the historical assumption of responsibility for that mechanism by the individual: that these individuals consider the outcomes of a particular mechanism as something they can be praised or blamed for.⁶⁹ An agent is moderately reasons-responsive (in a particular context) if (in that context) there is at least one possible scenario in which if it occurred and presented him with a reason to act differently he would act differently based on that reason for so acting. Further, he should recognise this reason as sufficient for him acting as he does. So in a situation where I bought a piece of second-hand vinyl, I am moderately reasons responsive if, had the circumstances been different - for example, had I reason to believe that it was overpriced - I might have acted differently, i.e. not bought the record, or tried to get the price reduced. For our purposes here we can consider an individual as morally responsible for an event if it is brought about in a controlled way, i.e. by one of his moderately reasons-responsive mechanisms.

4.1.2 Preparation as an Action

In the last chapter we considered whether IBPS can be seen as a type of intentional action. We concluded, as a consequence of the difficulty highlighted by the access puzzle, that because the content of the intention cannot match the new content brought about through the process of IBPS, the bringing about of a particular product of IBPS cannot be understood as intentional action (section 3.4 and subsection, 3.2.6). There are, however, three ways in which parts of the process of IBPS can be seen as an action. Firstly, on Strawson's (2003) account it can be seen as involving prompting and shepherding events within preparation (subsection 3.2.1). Secondly, we have suggested that under GCT an account of IBPS that utilises coarse-grained intentions such as 'to stop thinking about a problem' might contain trying events that are restricted to preparation, with the further stages of IBPS being seen as ballistic or non-voluntary parts of the action (section 3.4). Finally, under the advanced causal theory (ignoring the obvious difficulties) we can understand IBPS as intentional action guided by an intention-embedded plan to stop thinking about a problem now, but continue thinking about it later (section 3.4).

It can be argued that, even if some mental events cannot be understood as intentional action, an account of agency that depends on understanding part of these events as intentional actions will be sufficient for providing an account of moral responsibility for the consequences of those mental events in virtue of them being consequences of intentional actions. Understanding the relation of mental actions to

⁶⁹ I will not discuss this in detail; for more on this topic see Fischer and Ravizza (1998).

their consequences might then be sufficient to understand the agentive nature of many mental events, including IBPS.

4.1.3 Moral Responsibility for Action Consequences

This chapter will focus on the question of whether a successful account of moral responsibility for the process of the production of specific IBPS products is possible. That is, whether we can be morally responsible for the products of IBPS because of actions within preparation. We can begin with another example. You spend Saturday morning cleaning the front garden, hosing down all the paths. When the postman drops off the afternoon mail he slips on the wet paving stones. Are you morally responsible for the postman falling? You watered the path, so you are (a candidate for being) causally responsible for the postman falling. You did so in a controlled way, so you are a candidate for being morally responsible for washing the paths. The condition of force will apply to your action. If you were forced to wash the paths, then you would not be held morally responsible for watering the paths, and so you would not be morally responsible for the postman falling, but in this example you are not under any coercion. What about foreseeability? It seems here that dependent on what you know about the circumstances, you may or may not be morally responsible. If you knew the paths were slippery when wet, and you knew the postman was due before the sun dried the paths, then we will be more inclined to hold you more morally responsible than if you don't think the paths are slippery when wet or consider post (or another visitor) at this time on a Saturday quite unlikely.⁷⁰ The condition of foreseeability, then, will be central to any claim of moral responsibility for action consequences. I want in a moment to begin to consider whether individuals are responsible for the consequences of their actions in the preparation stage of IBPS. I will, however, begin with a related but distinct point: whether events brought about as part of an intentional action can themselves be considered intentional. I will begin with Bratman (1984). Analysing his account will lead us into a general discussion about responsibility for consequences. In the final section I will show why Stokes' account of minimal creative agency is not sufficient to support a claim of moral responsibility for the new content brought about by IBPS, although it may provide an account of moral responsibility for the general process of IBPS.

4.2 Bratman: Intending to Act and Acting Intentionally

Bratman's account of intentional action tries to widen the scope of the events that can be described as intentional. He suggests that some consequences of intentional action can be done intentionally even if they do not form part of intentional actions. Bratman (1984) puts forward a view distinct from the Davidson/Mele conception of intentions we discussed in the last chapter (section 3.1 and 3.3). Bratman agrees with Davidson and Mele that intentions are mental states (subsection 3.1.1). His view can be seen

⁷⁰ These considerations are distinct from any consideration of you intentionally watering the paths with the intention of making the postman slip.

as a type of causal view: intentions play a causal role, but he does not think they are reducible to belief–desire pairs, for he thinks intentions play an important role in the planning of our behaviour that is not captured by beliefs and desires (Bratman, 1984). His motivation for putting forward this view is a desire to understand how our intentions form part of our plans, and how this helps guide our behaviour. He is further motivated by a need to provide an account of action that is not limited by the condition of close fit required by the simple view (Section 3.2) and so can accommodate a larger body of events as intentional. I will not defend Mele’s view against Bratman’s view; nor will I contest any of the above points, because I am primarily interested here in understanding whether under these accounts IBPS will count as an intentional action. What I will do, then, is examine whether Bratman’s view will have any different consequences for mental actions than the account of action we discussed in the last chapter. This is important as Bratman’s (1984) account is motivated by different reasons e.g. that actors should be held morally responsible for some consequences of their actions even if those consequences are not a true part of the intentional action. We might therefore be able to include a different range of behaviour as intentional than considered possible under the accounts we explored in chapter 3. This might extend to the process and products of IBPS.

Bratman (1984) makes a distinction between what it means to intend to act and what it means to act intentionally. For Bratman, when an individual *A* intends intentionally, they intend something, but they need not necessarily intend to *A*. This prompts an obvious question: what is the connection between an individual’s intention and the fact that they intentionally do something? Bratman (1984) utilises the term *motivational potential* of an intention to make clear what type of acts can be seen as intentional under particular intentions. He suggests that *A* will be part of my ‘motivational potential’ for my intention to *B* if it is possible for me to *A* while executing my intention to *B*.

If I intend to *try* to hit a target with an arrow (*B*), my hitting the target with an arrow will be part of my motivational potential (*A*) for my *trying* to hit the target with an arrow, and so we may say that in executing my intention to *try* to hit the target with an arrow I intentionally hit the target with arrow. Here Bratman (1984) suggests that this relation does not show why it is the case that my hitting the target is intentional, even although I only possess the intention to *try* to hit the target; this will require further justification. What conditions will be required in order for my hitting the target to be intentional? Bratman (1984) suggests the following necessary conditions:

You intentionally hit the target with an arrow if

- (1) You want to hit the target, and so because you want to, you intend to try to hit the target.
- (2) You hit the target in the course of executing your intention to try to hit the target.
- (3) You hit the target in the way you were trying to hit the target.
- (4) Executing (2) and (3) depend on your relevant target-hitting skills.

Here we can see that conditions (1) and (2) set it up so that you have an intention to try, and that you actually achieve what you are trying. Conditions (2) and (4) begin to set the relation between your intention to try and a particular event being intentional.⁷¹

4.2.1 Two Extensions by Means–End Reasoning

By placing motivating potential as a separate entity from intention Bratman (1984) allows for the possibility that I can *x* intentionally - by which we mean something like being part of a means to an end, or more specifically as being a realised part of my motivational potential without forming a specific intention to *x*. Here Bratman is influenced by the idea that we should be held morally responsible for some of the consequences of our actions even if those consequences are not part of an intentional actions we do. If I blow up a tank, I might be morally responsible for killing the occupants even if I did not intend to kill them. Doing something intentionally, then, is to be morally responsible for that event. By ascribing moral responsibility on this basis the extent of behaviour we can be morally responsible for and that (for Bratman, 1984) we should see as intentional is extended beyond those behaviours that we specifically represent in our intention. This will include specific types of action consequences that can be understood as part of a motivational potential (that qualify under 1-4 above) but not part of the intentional action proper. This type of account has several advantages.

The first is already familiar (section 3.3): it addresses the problem of how to accommodate individual parts of actions as intentional without an individual being burdened with having to make specific intentions to do each part. Bratman (1984) suggests that what is required for part of an action to be seen as intentional is that it is part of an individual's motivating potential - i.e. follows (2-4) above - and for the agent to believe that the part of the action must be necessary in order to complete his intention. This achieves the same kind of result as Searle's account (1979, subsection, 3.3.1) or Mele and Moser's (1994, subsections 3.3.3 and 3.3.4) account of casual intentional actions. Conditions 1-4 above are very

⁷¹ We might worry here (as Bratman does) that there is no difference between intending to *A* and having an intention whose ongoing role includes the motivation of intentionally *A*-ing. Bratman (1984) overcomes this worry by considering the role of intentions in planning. If intentions are to take up roles in motivating and guiding our plans then they must be rationally consistent. If I have, as I suppose is generally the case, two or more intentions, say to meet Dave and to meet Tillmann, then in order to hold these intentions I must think they are rationally compatible. (For example, I might think that I don't have to meet Tillmann at the time I've arranged to meet Dave.) The constraints that are imposed by specific intentions will be more than those imposed by non-specific intentions that include motivational potentials; motivational potentials will not constrain actions in the way intentions do. Under Bratman's view, I will therefore be able to hold two action intentions that are *not* in conflict but have motivational components that *are* in conflict; whereas under the simple view, including all these constraints in the intention proper prevents me from holding these two intentions, because they would be rationally inconsistent. This will allow me to intend to do two things whose motivational potentials are in conflict, and this may be beneficial in achieving my goals. (Bratman gives the example of trying to play to computer game simulations, trying to hit the target on each, but knowing that this is impossible because of the way the games are linked.)

similar to Mele and Moser's conditions of following an intention embedded plan (see subsection 3.3.4). I will then not dwell on this form of extension. It will not be particularly helpful for understanding the consequences of IBPS or memory recall as being done intentionally, because under the first requirement the content to be retrieved will have to be available to the agent. The second form of extension may be more helpful for mental action.

Here Bratman (1984) extends instances of intentionally doing to include consequences of intentional action. He gives the example of intentionally wearing down a pair of sneakers. If I intend to run a marathon and do so intentionally then I may also intentionally wear down my sneakers. Bratman suggest that in order for such consequences to be done intentionally, and so for me to be morally responsible for them, I will have to believe that my intentional action x will result in my doing y . Here we can be said to do y intentionally. Perhaps being motivated by the condition of foreseeability, Bratman thinks that this will be most clearly seen as a case of intentionally doing when (1) I believe my doing x intentionally entails y , but also (2) I consciously notice y when I'm doing x - I note that I am wearing down my sneakers as I am running - and (3) the event y has some independent importance for me outside the end of my intention - for instance, in his example, the trainers might be family heirlooms. These extra conditions are supposed to give further conditions under which we have a strong case for considering action consequences as intentional. Bratman (1984) does not see these as necessary conditions, as are (1) to (4) above, but as extra conditions which, if present, will strengthen the case for counting events as intentional. This type of extension might provide a useful account of intentionally performing mental events and provide an account of moral responsibility for the consequences of intentional mental actions. If we take Bratman's (1984) conditions as set out, will the products of IBPS or memory recall satisfy these extra conditions for extension?

4.2.2 Bratman's View and Mental Agency

In our discussions in the last chapter concerning the nature of memory recall and IBPS, the access puzzle was the main difficulty in providing an account of intentional action for these mental events. In order for mental events to be intentional mental actions we required that the content of the intention represented the end result of the possible action. This limited the number of mental events that can be understood as intentional actions (see section 3.4 and subsection 3.2.6) and therefore the behaviour that can be said to be done intentionally to a range of mental events that did not include the activities of memory recall or IBPS. Does Bratman's account change this picture? If it can be said that I intentionally hit the target when I intend to try to hit the target, then we can get around the 'tight fit' constraint of the simple view, but does this directly affect the access puzzle? Firstly, it will still be the case under Bratman's (1984) view that for x to be part of an intentional action I must intend specifically to x . However, it will not be the case that for me to x intentionally I will have to have a specific intention to x . I can x intentionally if I have a specific intention to y and x is part of my motivational potential to y . Let us take the example of memory recall. I

intend to recall seven animal names beginning with 'g'. Can this be said to be an intentional action? Well, I want to recall seven animal names, and I intend to try to recall seven animal names. If I recall seven animal names, in executing my intention to try to recall seven animal names, and in executing my intention to do so, I utilise relevant skills in recalling seven animal names beginning with 'g'. If I then recall those names in the way I was trying to recall seven animal names, then it can be said that I recalled the seven animal names intentionally. Of course, it's still not true that recalling a specific name, e.g. 'goat', will be an intentional action.

To make sense of this we need to make sense of what relevant skills are being used, and what can be meant by recalling names beginning with 'g' in the way I was trying to recall seven animal names. If we take relevant skills to refer to a relatively predictable mechanism of recall - i.e. one that I am familiar with and use frequently - and the way I was trying to remember seven animal names to refer to utilising this mechanism - rather than, say, someone else prompting me - then most of my recalling will satisfy Bratman's (1984) conditions. I can then be said to intentionally recall seven animal names beginning with 'g' and so can be held morally responsible for doing so on his account. This result may not be counter to Mele's claim (subsection 3.3.5) that recalling seven animal names is not an intentional action but trying to recall them is. But it adds the further claim that the recalling is done intentionally, and so on Bratman's account we can be held morally responsible for our recalling of seven animal names.

But this does not help with establishing an account of moral responsibility for the bringing about of individual content (in the memory recall example, a specific name beginning with 'g'). Here the access puzzle prevents the remembering of particular animal names, e.g. goat, as an intentional action; but can Bratman's (1984) account make it the case that remembering each animal name is done intentionally? In order for this to be the case we would have to see the recalling of names as consequences, perhaps, of the act of trying to recall seven animal names. We have just examined the conditions under which Bratman suggests such side effects will be intentional. We should believe that they are necessarily part of the action we intend, we should be conscious of them occurring, and they must have some independent relevance for us. Let's leave this last condition aside for the moment; we will be able to construe cases where this is so. In our example of recalling seven animal names, it is not clear that remembering any particular name beginning with 'g' will be necessary for the completion of the action, as there are more than seven animal names beginning with 'g'. We will, however, be able to construe cases where recalling particular contents is necessary for completing the action, for example, of recalling John Lennon's birthday. It might be necessary to complete this action that we recall a specific piece of content, i.e. 09/10/1940. Even so, it is not clear that we will be able to be conscious of this event in the same way we are conscious of wearing down the shoes. We will be aware of the content coming to mind, but we will not be aware of the process through which it comes about in the same way the runner is conscious of wearing down his shoes.

This is a problem of access, and it raises the question of why we should think Bratman's (1984) conditions should apply to accounts of moral responsibility. What is his justification for including them?

Is it correct to say that we can only be morally responsible for action consequences we are aware of? We will return to this point after we examine whether Bratman's conditions can be applied to IBPS. For it is worth our time to show that neither IBPS products nor the general problem-solving process of IBPS can be seen as intentional under Bratman's account.

In the case of IBPS it seems that again we will have difficulty in suggesting that a particular content that arises is necessary to complete the intentional action. There will be many cases where different contents can be seen as a solution to our problem. If this is the case, then on Bratman's (1984) account no particular content can be seen as being brought about intentionally. Even if this were not the case in order for a side effect to be done intentionally (for Bratman), we would be required to hold a belief that bringing about the specific content is an essential part of the action. This brings us back to a version of the access puzzle, because in IBPS the individual problem solver does not know what that content is. It seem impossible for an individual engaged in IBPS to hold a belief that bringing about that particular content will be essential for his completing of that intentional action.

The condition of conscious awareness is even more problematic for IBPS than it is for memory recall. In IBPS no conscious awareness of incubation is available, and so we should think (if we agree with Bratman) that the bringing about of particular content cannot therefore be done intentionally. This may also rule out seeing the bringing about of a coarse-grained description of IBPS such as solving your problem as intentional. Although it might be the case that I can try to problem solve and in the course of my trying successfully problem solve, the condition of being consciously aware of doing this may rule out our problem solving intentionally and so rule out our being morally responsible for coming up with the solution on Bratman's (1984) account.

To summarise, it is difficult for Bratman's (1984) account to accommodate all instances of memory recall. This is because of the condition set by Bratman that in order for a side effect to be brought about intentionally we require a belief that such an event is essential to the completion of an action. This condition makes it difficult to accommodate the specific content of memory recall as intentional unless that content is the only content that will complete the action. Considering IBPS, the difficulties associated with the access puzzle will mean that it is impossible for the products of IBPS to be intentional. We will not possess the content that Bratman requires to be present in our belief about the necessary conditions of completing the action. The further conditions put forward by Batman, i.e. conscious awareness of the process, may make even intentionally problem solving difficult, although an account of memory recall may be forthcoming. As I noted, Bratman's conditions are quite exclusive and it is not always clear what motivates him to include them. There are many more accounts of moral responsibility for consequences, and it is worth our time to examine the central component of many of them, the condition that an action result be foreseeable, and ask whether an account of IBPS can accommodate such a condition.

4.3 Responsibility for Foreseeable Consequences

Bratman (1984) pursues the idea that consequences for actions can be intentional even if they are not intentional actions. He provides detailed criteria under which such consequences are intentional. As we have seen, IBPS will not be intentional under his account. A supporter of this line of argument might feel that the primary conditions Bratman suggests for an intentional consequence being necessarily part of an action are too strong. In our original example of responsibility for consequences we considered the question of whether an agent who washed the paths of his garden is responsible for the postman slipping on the wet stone. Under Bratman's view, in order for this to be the case, the postman falling would have to be intentional on our part. Bratman's view (1984) is partially motivated by the idea that in acting, not having a desire to bring about a consequence of that action does not prevent us from being responsible for that action's consequences, if we understand the consequences. This is the question concerning the *phenomena of double effect* (see Mele and Sverdlik, 1996). If a terrorist wants (has an intention) to blow up an ammunition dump, a known consequence of which is the destruction of a nearby school, will the agent be morally responsible for the destruction of the school even if he does not desire it or directly intend to destroy it? The straightforward intuition is yes. The agent is morally responsible even without possessing that particular desire or intention.

Bratman suggests that in some circumstances (see section 4.2) we can be morally responsible for action results and therefore we should consider those results intentional on our part. Bratman (1984) quotes the utilitarian Sidgwick (1907), in suggesting that such ascriptions of whether an agent is intentional depend to some extent on whether we should be held morally responsible for the consequences to be ascribed. But, as Mele and Sverdlik (1996) suggest, whether the word intentional applies in these cases may not be important or relevant to an ascription of moral responsibility. That is, it is a linguistic debate. What is important is not the range of application of the linguistic term of doing intentionally but the conditions under which we should be held morally responsible for the consequences of our actions.

Which conditions are relevant? The early utilitarians, Bentham (1789), Austin (1863), and Sidgwick (1907) all suggest a similar condition on moral responsibility for the consequences of our actions: that the consequences are foreseeable. Mele and Sverdlik (1996) suggest they were moved by considerations concerning double effect; that in circumstances when you do foresee a consequence of your action you cannot escape moral responsibility for the consequences simply because you do not desire them. As such, foreseeing a consequence of your forthcoming action makes you a candidate of praise or blame for those consequences if they occur.

Foreseeability, then, has been utilised widely as a general requirement for ascribing moral responsibility for action consequences. Foreseeability can be understood in different ways. One obvious way is to understand what is foreseeable as whatever is immediately accessible to the agent i.e. information that he is in possession of at the time of action. For example in throwing a dart at a dartboard an agent can foreseeably hit any number, or the bull's eye, or the man sitting under the board, if he is

aware of these things at the time he throws the dart. He might therefore be morally responsible for hitting the bull's eye even if he only intended to hit the board; equally he could be morally responsible for hitting the board even if he intends to hit the man. Here foreseeability is limited to immediately accessible possibilities. Foreseeability might also be extendable beyond these cases. Take this example: you are attempting to shoot the King. Sitting beside the King is the Queen, whom you don't register. Even if you don't register that the Queen is there, you might be morally responsible for hitting her, because it might be expected that you *should* have checked properly whether others were around. Here, deliberating about the matter may bring to mind a range of further possible consequences that an agent might be expected to take note of.

We can apply the condition of foreseeability as a singular requirement for moral responsibility of action consequences to the specific products of IBPS. Even if parts of preparation are understood as intentional actions, foreseeability itself - in both its limited and extended guises - will exclude an agent being morally responsible for any specific piece of content brought about through IBPS, because, as we have suggested, such foreseeable consequences will not solve an individual's problem. Agents here are thought to have engaged unsuccessfully in such deliberations in preparation, and it is the unsuccessful nature of such deliberation that leads to incubation. Products of IBPS will therefore fall outside both notions of foreseeability, for they are in practice not accessible by deliberation. It may be possible for an agent to be morally responsible for problem solving as a consequence of actions done in preparation, because the event of successfully problem solving is foreseeable. In the final section I will examine this possibility for Stokes' minimal account of creative agency.

4.4 Minimal Creative Thought: A Non-Accidental Agency Clause for Creativity

Stokes (2011) suggests that one way of separating events that are agentive and therefore possible candidates for being creative from non agentive events is to consider that agency depends on events being non-accidental. He puts forward an account of minimal creativity where for a thought to be creative it is sufficient that it is brought about in a non-accidental way; it is p-novel, and it must not have been tokened before the time it is actual tokened. Here we are only interested in Stokes' agency clause.⁷² He motivates his agency clause by comparing events brought about by an agent and those that occur without an agent. He suggests that cloud shapes or other natural phenomena cannot be considered creative even if they are novel. On the other hand, man-made objects, pictures, paintings, etc., are candidates for the ascription of creativity, and this might be the case despite the fact that they exhibit qualities similar to natural products. He suggests that only objects that have been generated through agentive behaviour i.e. are non-accidental, are candidates for the ascription of creativity. This of course, as Stokes (2011) notes, does not mean that mental events we generally think of as creative are necessarily agentive. They have to be of the right

⁷² See chapter 1, section 1.4 for reasons why such definitions of creativity are problematic.

character. If Stokes is right about the concept of creativity, then it may be that none of our mental events are truly creative.

Stokes (2011) agrees with some of our considerations about IBPS and intentional action: that IBPS is not an intentional action, and nor are the specific contents brought about by IBPS part of an intentional action (section 3.7). But he suggests that some mental events, if they meet the above conditions, can be seen as creative. This will entail that they are agentive at least in as much as we can be morally responsible for them. To be morally responsible for something, in the sense Stokes has in mind, is to be an appropriate candidate for praise or blame for the occurrence of a particular event. One way of understanding this is to suggest that, if part of creative thinking is an action, then we might be morally responsible not only for the action but also for the downstream effects of that action.

Stokes (2011) is talking generally about events that might be considered creative. He is not talking specifically about IBPS. He does, however, describe a general category of what he calls flash phenomenology that is likely to include IBPS. He suggests that even if these events are unusual the overall operation of producing creative products, even if they include flash phenomenology may be agentive if we can be considered morally responsible for producing those products. Moral responsibility here rests on the fact that responsibility can be traced back to events of deliberative cognition. In the case of IBPS, we might consider that actions in preparation, such as the acquisition of information, or prompting the mind to reconsider some possibility, etc. (section, 4.1.3), occupy this role of agentive activity, and that we can trace a non-accidental link to future consequences that we can be held morally responsible for. We would then be morally responsible for creative products produced in revelation, as we have described it, because of the deliberate mental event, perhaps the intentional actions, in preparation. As we have considered, not all consequences will be connected in the right way to previous actions for us to be morally responsible for them. Stokes suggests that a minimal condition of consequences being connected to action in the right way is that those consequences be non-accidental.

In light of the difficulties we have encountered in describing the whole process of IBPS as an intentional action, Stokes' (2011) approach seems at first glance to be a promising avenue of investigation. It leads us back to some considerations already under discussion. For instance, how exactly can preparation be seen to be causally efficient in the right way for all the stages IBPS to be seen as agentive (subsections, 3.2.5 and 3.4.1)? There are two considerations here that must be separated: the conditions for seeing the products of IBPS as intentional actions will be more stringent than the conditions for seeing the agent as morally responsible for the products of IBPS that arise as a consequence of action within preparation. We will explore Stokes' claim that an agent can be morally responsible for the products of IBPS in the next section. First, we will consider whether Stokes is right to think that creative acts with a character similar to IBPS can be seen as agentive, even if they are not part of intentional actions and what such agency would entail.

Stokes (2011) seems to be aware of the access puzzle (section, 3.2.3). He suggests that for an account of intentional creativity we will require a criterion that is less stringent than a one-to-one match of intention and action result, but stringent enough to exclude chance interventions. To back up his case he adopts a Davidsonian account of intentional action (sections 3.1, and subsection 3.1.2). First he puts forward this weaker claim: even if creative events such as IBPS are not intentional actions, they are nevertheless agentive. Take the case that Hamlet kills the man behind the arras. He intends to do so and can be said, if he meets Davidson's (1963) conditions, to intentionally kill the man behind the arras. Unknown to Hamlet, the man behind the arras is Polonius, but Hamlet cannot be said to intentionally kill Polonius under the description that he intended to kill the man behind the arras. But even so, we do not deny that Hamlet did in fact kill Polonius. He just did not do it intentionally. Here we can see that there is a minimal sense in which Hamlet, through acting, killed Polonius. This is the case in virtue of the different descriptions that can be attributed to the event. Here I think the claim amounts to no more than Hamlet's acting entered the causal chain, the result of which causes Polonius's death: to use the terminology we utilised above, Hamlet can be said to be causally responsible for killing Polonius. Hamlet is the cause of Polonius's death, and this is not conditional on Hamlet having the intention to kill Polonius. The sense of agency involved here, if there is one, is not secured by an agent acting in accordance with or being sensitive to reasons: we can say just as well that the wind is causally potent in bringing about your tree falling down. This is because the condition of foreseeability excludes Hamlet from being morally responsible for killing Polonius. The fact that Hamlet did not believe Polonius was behind the arras means that he cannot be motivated to act based on his reasons for wanting to kill Polonius (even if he had such reasons). It seems that a necessary condition of agency will be that an agent is causally responsible for bringing about an event but we should not consider it a necessary condition of moral responsibility. As we discussed above this will require the common condition that the behaviour is brought about in a reason sensitive way.

A stronger claim could be made that Hamlet is morally responsible for Polonius's death. The sense of agency involved here is dependent on there being an intentional action that is connected to the killing in the right way - i.e. the killing of the man behind the arras that Hamlet did, and did intentionally - and on the fact that the consequence of that action is that Polonius was killed. The conditions under which an individual is morally responsible for action consequences are not straightforward. Will Stokes' criterion of an event being a non accidental result of an intentional action be sufficient for ascribing moral responsibility for the products of processes like IBPS because they are the result of prior intentional reasoned activity on our part?

4.4.1 Minimal Creativity and Responsibility for Creative Products

In Stokes' (2011) account of minimal creative agency, he suggests that acts of revelation might be agentive because they have been brought about through earlier reasoned behaviour. He suggests that the

consequence of these prior actions can be considered agentive if they are not accidental. Stokes elaborates what he means by non-accidental in the following example. I intend to lift opened, heavy paint pots across a room; I am not morally responsible for the consequences of splattering of the paint or the splattering of the paint in an aesthetically pleasing way because these were not intentional actions on my part. The responsibility associated with the aesthetic splattering Stokes suggests is trivial, for it is only loosely connected to my intentions and therefore accidental. This example is useful here because it is easily converted to IBPS. In a case of IBPS, if we consider a case where an intentional action taking place in preparation is equivalent to moving the paint, and the paint being spilled in an artistic way is equivalent to the exact content that is brought about. In this sense the exact content is accidental, and so we are only trivially responsible for it. It seems the exact content here should be considered accidental as it is not represented in our intention and therefore such an account will not allow for a claim of moral responsibility for the specific content brought about in revelation. However it might be possible to gain a claim of moral responsibility for the problem solving; this will depend on us considering the problem solving as at least being a non accidental consequence of our intentional action.

In Stokes' account of agency for consequences what it means to be non-accidental is meant to be vague. It is meant to be a place holder for a more complete account of what is required for moral responsibility. We might be morally responsible for problem solving if it is non accidental and the result of our intention to problem solve. It may be then that if we have an intention to problem solve and this non-accidentally results in our problem solving, even if this is not an intentional action we will still be morally responsible for problem solving.

Further, any condition of foreseeability has a close match to a condition of being non-accidental. There may not be an exact match up here but the range of possibilities seems intuitively quite close and this might be the case because being an accident might entail that at least an individual involved did not think a particular event likely to happen.

Including both will not prohibit us being morally responsible for problem solving for given the right epistemic background that a foreseeable and non -accidental consequence of your trying to problem solve or stopping thinking about a problem will be the finding an answer to your problem. It seems then that an account of moral responsibility for consequences will be applicable to the consequences of an intentional action (such as those that can take place in preparation (subsection, 4.1.3)) within IBPS but no account of moral responsibility will be available for the specific products of IBPS. Rather accounts of moral responsibility will be limited to an individual being responsible for the solving of the problem under consideration. It is worth our time examining the nature and merits as well as the difficulties of an account of moral responsibility for such a process.

4.5 The Minimal Agency Model for IBPS

The best agentive model for IBPS we have encountered so far can be put like this. We act intentionally within preparation, in our searching for new information, in possible prompting actions, and in possible focusing actions. We can have intentions to do these things, and they can properly be described as intentional mental actions. These actions have causal consequences, and it is probable to think that such actions are important causal precursors for insight. Since solving your problem is a foreseeable consequence of these actions, then we can be morally responsible for solving these problems through insight. I will call this the *minimal agency model for IBPS*. The picture here that comes from Stokes (2011) and involves Strawson-type prompting and shepherding actions (subsection 3.2.1) in preparation is not without merit, but it is important to get clear the conditions under which such a view is applicable.

Under the minimal agency model of creativity, the actions involved are based in preparation; those that are most clearly actions, the prompting of thought or the shepherding of the mind to concentrate, are relatively disconnected from the consequences of solving the problem through insight. The insightful understanding gained in revelation, although causally connected to events in preparation, may only loosely be connected to any particular action. Consider the act of focusing on a particular aspect of the question. This might be an action, but if that particular aspect of the question is not useful in providing an answer, the connection between the action and the subsequent revelation is very loose.

In the case of IBPS, many of the actions in preparation may have had no effect on the problem solving process at all. The best picture here, if you like, is a ‘machine gun’ account: lots of activity we consider actions in preparation and the downstream consequence of coming up with a solution. Some or part of this activity will be partial causes of the resulting insight, but the causal connection between them will remain unclear. We will never be certain which of these events actually resulted in you solving the problem. But nevertheless, we may think that agents of those actions can be morally responsible for the collateral effect of those actions, including revelation. Just as in the case where you sprayed a belt of ammunition at oncoming troops, you would be morally responsible for the deaths of the soldiers killed from the bullets fired from your machine gun even if you could not distinguish which burst of the trigger issued the bullets that killed the individual soldiers.

The condition of foreseeability still places some constraint on which episodes of IBPS can be considered agentive. It will have to be the case that in acting as they do in preparation, the individual concerned must consider that it is an open possibility that his so acting will result in answering the problem at hand. This might seem quite straightforward, but the more difficult the problem is, the less confidence an individual may have that the activity he is engaged in will result in an answer to the problem being produced. However, in most circumstances a reasonable familiarity with problem solving will make it the case that solving a problem is a foreseeable consequence of actions in preparation.

4.6 Moral Responsibility for the Specific Products of IBPS

An important limitation of the minimal account is its inability to ascribe moral responsibility to specific products of IBPS. The fact that accounts of intentional action in many situations cannot provide a claim of agency for the particular content is particularly problematic for IBPS. In IBPS we do not just want to say that the process or part of the process is agentic, but that the bringing about of specific products is an agentic mental activity, one that forms part of the sphere of mental agency. If we want to be able to understand particular ideas as belonging to individuals, and individuals as morally responsible for those ideas, it is important that the bringing about of individual content comes within that sphere - that the activity of bringing about specific content can be considered part of the personal activity of the agent. For now it is worth our while suggesting why such a limitation is important.

We can begin to motivate our concern by considering the case of forming beliefs. Consider the case when I form the belief that all killing is wrong and that therefore I should withstand the draft. An account of moral responsibility for forming a non-specific belief will not secure our moral responsibility for the actual belief that we form. Here, me being morally responsible for making up my mind in a non-specific way does not entail that I will be morally responsible for the particular belief that I form, i.e. that I should not kill under any circumstance. Being morally responsible for the specific belief entails far more than being responsible for the forming of a belief on a given topic. An account of moral responsibility for making up my mind does not entail that I am blameable for the exact choice I made, but only for the actual fact that I made a decision.

A similar type of concern is present in IBPS. In IBPS we are about the business of producing new content. One of the important aspects of IBPS is that it can often generate novel and useful products, products that we have never thought of before or even considered as possible solutions. Einstein's development of relativity is not important just because it solved a problem in physics, but because it gave us a whole new way of looking at the universe: a more straightforward answer (perhaps in this case less satisfactory) would not have resulted in the revelation in physics that resulted from Einstein's theory. It seems that we want to be able to hold Einstein accountable for the exact theory we came up with not just that he came up with a solution to a particular problem. It seems intuitive, then, that a complete account of agency for IBPS will include an account of how we can be agentic in producing the exact product that we do produce. Without such an account, any theory of agency for IBPS will be at best incomplete. This becomes apparent when we consider the novel aspect of IBPS products. As we discussed in subsection 1.6.2 IBPS will often produce novel solutions: solutions that are not foreseeable (section, 4.3). It is this aspect of IBPS products that is the primary indicator of whether they are creative or not. But in the case of minimal creative agency no claim of moral responsibility can be made for the content of the product. And since the novelty is contained within the content no claim of moral responsibility can be made for the novelty inherent in the content. It seems here we will not be morally responsible for the novel aspect of IBPS product.

A further concern rests on the fact that moral responsibility, here, is based on a consequential account. It is intuitive that we are responsible for consequences, but in many cases - especially cases where the causal connection is distant and involves a degree of chance - the fact that you can be considered morally responsible will not always dictate as strong a claim of moral responsibility as intended action results. Here, although you may be blameable or praiseworthy, the strength of that blame or praise will be reduced by the circumstances of the episode in question. Here we can consider that moral responsibility is not an all-or-nothing condition, inasmuch as the judgments about whether you are blameworthy or not will often be qualified. Consider a case of killing. Killing by driving dangerously is quite different from killing by intentionally running someone over. Here, the second, as a consequence of a direct action that is an episode of reason controlled behaviour, is viewed as much worse because the individual running someone over can be seen as part of his reasoned behaviour. In the case of IBPS, the individual can be morally responsible for the problem being solved, but he can't be morally responsible for the problem being solved intentionally. In such cases, the agency involved is of an indirect kind. It is subject to ascription of praise or blame, but it is not based on a direct episode of controlled behaviour that is associated with an action but rather a foreseeable consequence of that action. We will return to this point again in chapter 5 (subsection 5.7.2) after I have developed a positive account of agency for the specific products of IBPS.

It may be, then, that minimal agency for IBPS based on the consequences of actions within preparation is possible. My intuition here is that the minimal account is not particularly satisfactory. No account of moral responsibility will be applicable to the particular content of IBPS. This may be particularly important for IBPS, for we are not just concerned about the activity and agency involved in bringing about a solution, but the agency and activity involved in bringing about a *particular* solution. In any ascription of creativity we are interested in the content of that solution. Accounts of moral responsibility for IBPS provide only a minimal sense in which we are agentive in some circumstances. It is my contention that this account fails to capture a full understanding of the personal activity involved in IBPS. At the very least these consideration provide us with a reason to look for a more satisfactory account. Providing a deeper understanding of IBPS that captures the full nature of the activity and agency involved will be the work of the next chapter.

4.7 Conclusion

In looking at responsibility for action consequences, we have completed our exploration of whether the process of IBPS can be seen as an intentional action or whether products of IBPS can be seen as consequences of actions that take place in the preparation stage of IBPS. No account of IBPS as an intentional action or as something done intentionally under Bratman's (1984) account was forthcoming. In looking at action consequences, no account that provided responsibility for particular products of IBPS

was available to the agent. An account of minimal creative agency that allowed for moral responsibility for solving the problem was available.

Chapter 5

The Activity of Insight

5.0 Introduction

In the last chapters we saw the deficiencies inherent in conceiving the activity of the mind as a type of mental action. The analogy we drew with physical action was unable to capture the mental activity of new content coming to mind and since the activity of bringing about new content is essential for IBPS, in resolving the cognitive problem faced by the agent, we were unable to capture the active nature of IBPS. If we want an account of agency for IBPS that captures the role of the individual in constituting the content of the mind we need a fresh perspective.

In this chapter I propose to take a leaf from Richard Moran's (2001) *Authority and Estrangement*. Moran's key observation in that work is that when we occupy a particular first person stance in forming some of our beliefs, we are in an active position of deciding on, and participating in, bringing about or constituting those mental states. For Moran, when we deliberate on what we shall do, for instance, to make up for our absence at a friend's party, the arriving at a particular belief is fundamentally our activity. It is the individual who decides and can be held responsible for so deciding. Moran calls this particular first person perspective the *deliberative stance*. As we shall see it is quite distinct from the intentional activity commonly associated with agency that we explored in chapter 3 and 4.

For Moran (2001) understanding the first person perspective is essential for understanding the sense in which deliberation is active. It is my conjecture in this chapter that adopting Moran's deliberative account of agency, an account that depends on understanding the role that the first person perspective plays in constructing new content, can provide a beneficial way of understanding the agentic nature of the process of IBPS. Within this chapter I will explain Moran's account of why the activity of judging from a first person perspective is the activity of an agent. My mission here is not to defend his view against all possible objections; rather in explaining it I hope to provide reasons for believing it and show the benefits that such an account has for our understanding of IBPS.

In order to provide a new account of the activity of IBPS we will need to come to understand the particular active nature of Moran's first person deliberative stance. I will then be in a position to argue that within revelation, a key stage of IBPS, we occupy the deliberative stance. If the creation of new content within revelation occurs from that stance it should be considered to exhibit the activity and agency that the deliberative stance affords. Furthermore, the nature of revelation - devoid as it is of any intentional activity- provides, to date, a unique window into the type of deliberative activity Moran has in mind, and in doing so contributes to the debate about the nature of that deliberative activity.

So it is my contention that the accounts of mental action we investigated in chapters 3 and 4 fail to capture the type of activity involved in IBPS. Much work will have to be done in delineating exactly the nature of the activity Moran thinks is important for grounding agency. In particular, we need to separate the deliberative activity Moran (2001) is concerned with, in deliberation and the formation of beliefs, from

the more familiar type of intentional activity we were concerned with in chapters 3 and 4. So, I will first turn to considering the nature of deliberative judgments and the constraints on control that are constitutive of that nature.

5.1 The Voluntary Nature of Mental States

A familiar take on the voluntary nature of our mental life suggests that some aspects of the mind are more free than others. Imagining and desiring are less constrained than intentions and actions, and judging and believing may be even more tightly constrained⁷³. Take as an example, Hieronymi's (2009) framing of the *toxin puzzle*⁷⁴. The toxin puzzle is supposed to show that the formation of intentions is constrained in a different way to actions. In Hieronymi's (2009) example, a psychologist asks you to form an intention about two different actions: firstly to form an intention to drink from a glass of water and secondly (and separately) to form an intention to jump from a third storey window. In both cases you are offered a reward to form the relevant intention. To gain the reward you do not have to act, only form an intention to act. In the second case Hieronymi (2009) thinks that it is impossible to form the intention to jump from the window because it is not under your control to form an intention on its own without committing to the act your intention is about. What is required for you to form an intention to jump is a reason that makes the act worthwhile, not a reason that makes the intending worthwhile. Here there is a distinction between the control you have over an intentional action, i.e., that you can be moved to action by any reason that makes the act worthwhile, and the control you have over forming an intention, where it is only possible to form an intention that is motivated by a reason to perform the act that you wish to intend (Hieronymi, 2009). This is brought home by Hieronymi's explanation of the first case, where the reason you can intend to drink the water is the reward being sufficient reason to drink, and the reason you cannot form the intention to jump out of the window is the reward not being sufficient reason to jump out of the window. Intendings, then, cannot be controlled in the same way intentional actions are. We might think that this is a necessary feature of intendings that they can only be motivated by reasons that pertain to the action they

⁷³ For the voluntary nature of imagination see Strawson (2003), Dorsch (2009); for intentions see Kavka (1983), Mele (1992a); and for the formation of belief, see Moran (2001), Hieronymi (2008, 2009), Smith (2005).

⁷⁴ Kavka's (1983) original puzzle is slightly different. The example involves intending to drink a potion tomorrow morning that will make you ill for a day. If you do so you will get 1 million dollars at 12 o'clock tonight. The puzzle involves the fact that you cannot form an intention to drink the poison because you are aware of the fact that you will already be in receipt of the money and therefore your motivation for drinking the poison tomorrow morning will have gone. This awareness makes it the case that you can't form an intention to drink the poison tomorrow morning. The possible motivation of gaining the million dollars will be insufficient because you will gain receipt of the money before you have to drink the poison. Hieronymi's (2009) case is different in its set up but the important point remains between the two. Both are suggesting that you cannot form an intention to act without sufficient reasons to act in that way, and no reason to so intend can bring about the formation of an intention to act unless it itself also motivates the action. For further discussion see Mele (1992c), Levy (2009).

represent; after all this is the function we expect them to perform. We may have the intuition that, for similar reasons, judging and reasoning and many other types of mental episode are also not controllable in the same way as actions are or indeed in the same way intentions are, and that the voluntary nature of different mental states will reflect the role they play in our lives.

In respect to belief formation it is a familiar story since Pascal (1656) that I can't just decide to believe in God because it is incumbent upon me by society, or because I have some reason for thinking such a belief will be valuable or profitable to me. Instead, at best, if I want to acquire such a belief I must act indirectly by reading the Bible or attending church, in hope that such actions might indirectly bring about the desired change in my beliefs. For philosophers a question arises as to why this is the case. Why are we not free to alter our mental states any way we choose, and why are these different types of mental states constrained in different ways? It seems likely to many that the constraints on different mental states tell us something important about the nature of those states. The difference between the voluntary nature of a belief and a desire to some extent makes those states what they are. I will turn now to examine the voluntary nature of judgements.

5.1.1 The Voluntary Nature of Judgements

The question of the voluntary nature of judgements has been widely discussed⁷⁵ and there is a range of opinion on just how voluntary judgements are. At one end of the spectrum, Strawson (2003) suggests that judgements are entirely non-voluntary, and that deciding on the outcome of deliberation is not a matter of action (he thinks mental actions are limited to priming and shepherding cases (subsection 3.2.1)). At the opposite end we might consider that we are free to judge in any way we choose. However most modern philosophers suggest a third way. They consider that judgements have unique constraints, to which actions are not subject, but they nevertheless still contain an agentic element (Moran, 2001; Hieronymi, 2008, 2010; Smith, 2005; McHugh, 2011; Toribio, 2011). The extent to which we should consider judgements as voluntary or agentic will depend on which of these accounts you find convincing.

McHugh (2011) presents us with an intuitive example. Take any issue which you care about, for example, that people everywhere should have the right to vote. He suggests that if someone offered you a hundred thousand pounds to make a judgment against the way the evidence points, you would be unable to do so. You would not be able to make a judgement based on reasons that are not aimed at the truth of the issue under consideration. You would of course be able to lie about what it was that you judged. But in making a judgment McHugh is, familiarly enough, suggesting that such reasons will not play a role in your process of judging. At the heart of the debate is a discussion about the control we can exhibit over judgements. But there is also a range of opinions over exactly what judgements are.

Judgements as we normally understand them are mental episodes or events that are often

⁷⁵ For an account of involuntarism, see Williams (1973) for discussion see Bennett (1990), Engel (2002), Owens (2003) James (1905), Helm (1994). For an experiential account of voluntariness, see Dorsch (2009).

reportable. That is, we have a conscious phenomenal experience that, at least in some cases, is one with episodes of judging (McHugh, 2011)⁷⁶. Judging can be seen as an acceptance of the truth of the consideration in question but it might also include the accumulating of evidence. Where exactly the mental episode of judging begins and ends is up for debate and I don't think it is necessary here to argue for a specific conception of what judgements are⁷⁷.

What is clear and common to all accounts is that judgements involve a specific event of commitment to, or endorsement of, a particular piece of content (McHugh, 2011; Toribio, 2011). For our purposes it will be enough to consider judgements as the endorsement of content, of a commitment to the truth of a particular proposition. This stage of commitment, as we shall see, is not intentionally under the agent's control. To come to a new understanding of the voluntary nature of judgement will occupy us for the rest of this section and sections 5.2 and 5.3. We can begin by eliminating a particular type of control that is not possible in judging: *reflective control*.

Reflective control suggests that it is possible upon reflection to decide, one way or another, what the product of your reasoning will be (Hieronymi, 2009). For example, I have a query about whether I should emigrate to New Zealand. In bringing to my mind this question, we might think that it is possible one way or another to decide (in the moment) whether or not to emigrate to New Zealand; that I can intentionally bring about one specific outcome over another. In this way, I would be seen to be agentive over the object of my reasoning. This is wrong. We do not have this type of control due to two kinds of considerations for the following reasons. Firstly, it seems that we cannot control exactly what considerations or attitudes are brought to bear upon the object of our reasoning; secondly, we do not have control over the result of the reasoning process (Hieronymi, 2009). We cannot intentionally decide *x* or commit to *x* where *x* is a specific result of rational processes. If we agree with Toribio (2011) no argument for this conclusion is, perhaps, necessary, for it seems when we openly reflect on the question our own experience leads us to conclude that such judging is not open to our arbitrary choice.

Hieronymi (2008) as well, suggests that an important difference between actions and judgements is the type of reasons that are involved in the agent's relation with the event. Hieronymi (2008) discusses the voluntary nature of deciding but her discussion suggests that deciding is equivalent to forming beliefs. She suggests that believing cannot be understood as an action because it entails its own kind of answerability that is distinct from the answerability that is possible in action. When you believe something you are answerable to the reasons that you take to indicate the truth of the belief. For example in judging that it is raining outside, the evidence that indicates this state of affairs and forms part of your reasoning for coming to believe it is raining is the only kind of explanation that accounts for your holding that belief. This is in contrast to the type of answerability that is involved in an action, where you are answerable for

⁷⁶ No argument rests on the fact that judgments are always reportable

⁷⁷ McHugh 2011 suggests that judgements should be identified solely with the event of commitment to the truth of a propositional attitude, see also Owens, 2000).

any reasons that you take to be in favour of you so acting. Hieronymi (2008) is pointing to the range and variety of the reasons that can support your acting, and those that can support your judging, deciding or coming to believe. This latter set of reasons is a much smaller subset that has to be supported by the particular evidence that aims to represent the world correctly. Whereas the reasons you have for acting are not so constrained (even although as we have seen in the toxin puzzle our reasons for intending have specific constraints of their own): in acting we are answerable for any reason that we take to motivate that action. These reasons are not constrained by being aimed at representing the world, they are reasons bound up with desires and other motivational forces that are not necessarily governed by any notion of correctness or truth; they can be directed at internal goals that are not constrained, as judgements are, by the truth of the proposition in question.

The stage of commitment, perhaps the only necessary stage of judging, is not then under our reflective control, it is rather under a normative constraint - it is operating under the constraint of being subject to a particular type of answerability, of being justified, supported and motivated only by reasons that conform to a particular normative standard of representing how things are. It is then not under our reflective control to judge x where x is the specific content of a judgement. It is perhaps under our control to judge y where y is a question that we want to judge e.g. judge whether a particular essay is a good one. Marking an essay is a clear example of intentional judging of, that is I can form an intention to judge the content of the essay, but it is not under our intentional control to specifically bring about the particular content of what that judgement will be.

In claiming that judgements are not under our reflective control, we are also claiming they are not under our intentional control. That is, we eliminate judging as a process that can be understood as an intentional action because our intention to bring about the specific result of our reasoning does not form part of our motivation for deciding. There are however cases where we intend to judge x and x does in fact come about. For example, you can have the intention to judge an essay as an A, perhaps because you unfairly favour a student or because you have always wanted to give an A but have as yet not had the opportunity. It might be, then, that on reading the paper and forming a judgment on its content, you judge the essay should receive an A grade. You come to the realization that the paper is an A and this occurs exactly through an episode of mental judging that is bound by the constraints we have just been discussing, that is based only on reasons that are consistent with correctly representing the world. But as we discussed in the last chapter (section 3.2), for an event to be an intentional action it is not enough that we have a match between what an intention represents and what actually occurs in the world: remember our trip to the NYPD. What would be required for judgements to be intentional actions is that the agent can intentionally bring about what the outcome of a judgement will be - that his intention is important in bringing about the result of the judgement. This is impossible for the range of reasons upon which you can form an intention will be distinct from the range of reasons that motivate your judgements. Although you may well be able to form an intention to judge x , in pursuing that intention and actually judging you are

constrained by the reasons in the world to which the judgement is sensitive and thus any motivation garnered from the intention is lost. You may be able to assert that this essay is an A, and do so intentionally by so pronouncing it to the world, but to judge as such requires you give up on any personal reasons and answer only to reasons that conform with representing the world correctly⁷⁸. The process of judging cannot then be an intentional action. It cannot be under our reflective control because it is not possible for our intentions to be the decider, to be causally efficacious in bringing about the content of a judgement. For that judging to remain a judgement it must conform to its own nature, as such coming to make a judgement is subject only to reasons that conform to representing the world correctly.

Judgements, then, should not be seen as voluntary in the same way actions are. This conclusion has led some philosophers, such as Strawson (2003), to argue that the mental events associated with judgements are mainly involuntary. In chapter 3 (section, 3.2.1) Strawson's account of mental action was presented. In his account mental actions were limited to prompting and shepherding actions that, crucially for Strawson, bring about no new content as part of the action. These types of intentional mental action might be present within the process of deliberating or judging. For example within the process of judging I might focus my mind so that I can concentrate on the deliberation I am trying to carry out, or I may prompt my mind to consider the implication of a particular point, that at that time I feel I have not fully explored. On Strawson's (2003) account mental action in judgements goes no further than this. Crucially for Strawson the central component of committing or endorsing the judgement that takes place should not be seen as the activity of the agent for it is not an intentional mental action. Under such an account, judgements are viewed in the main as involuntary events that happen to us, with particular mental actions limited to prompting or focusing the mind on those events rather than participating in them. Here I think we are not misrepresenting Strawson if we say that he thinks that judgements, excluding shepherding and priming actions, are both involuntary and have little to do with agency⁷⁹(Strawson, 2003).

But this picture of judgements as passive mental events has seemed to many to underestimate the active nature of judgements. Even if judgements are not intentional mental actions they may nevertheless be in some other sense the activity of the agent. This line of thinking has led some philosophers (e.g., Hieronymi, 2009; Smith, 2005) to suggest a new type of account of how our mental activity is controlled

⁷⁸ It may still be the case that your intention to judge x (e.g. an essay an A) is important in you actually making a judgement. Here we can suggest that the part of the content of your intention that is important in you making a judgement is your wanting to judge the essay, not your wanting to judge the essay an A. Even if your intention to judge an essay an A is a partial cause in you judging an essay an A. We can say that it is a cause in virtue of you intending to judge, and not because you intended to judge an A. So the specific outcome is not within your intentional control.

⁷⁹ Strawson (2003) gives a strong indication that the coming about of content is not a matter of agency. To be fair to Strawson he does allow that mental activity can be understood as an involuntary activity that is the result of an individual's reason. Strawson then allows room for agentive accounts based on reasons, in this broadly Kantian sense, rather than action. It may be that this suggestion is not in conflict with my own account of IBPS as deliberative. Strawson does not provide enough detail to judge whether my account is in conflict with the notion of reasoned involuntariness he mentions.

based on a distinction between passive and active mental episodes. These accounts are to some extent based on Moran's (2001) account of the personal activity involved with taking up the deliberative stance. Hieronymi's (2009) account is a theory of how our reasons control our mental lives. She uses the title of *evaluative control*. Where at heart, evaluative control is the ability to respond to reasons. All things being equal, if other reasons had been presented we would have acted in accordance with those reasons⁸⁰.

In this chapter I will suggest that evaluative control is present within the process of IBPS. It is hoped that adopting evaluative control will be beneficial for our understanding of IBPS. The justification for this endorsement does not rest on any decisive argument against other views of judgments but in the gains to be made by such an endorsement in our understanding of IBPS. In endorsing Hieronymi's (2009) and Moran's view (2001) I am not in direct contradiction with Strawson's (2003) claim that the bringing about of new content in judgements is not a matter of action. But I am in direct opposition to the claim that judgements are not an agentic activity. Further, when Strawson briefly discusses problem solving he hints heavily that he considers the bringing about of new content is not only involuntary but also not a matter of agency. I am in direct opposition to the claim that problem solving is always non-agentic, for later in this chapter I will argue that IBPS contains episodes of agentic activity that include the bringing about of new content. In order to propose that the activity associated with evaluative control is present within IBPS we need to understand the nature of evaluative control. It is to this new understanding of the voluntary nature of judging that I will now turn.

5.2 Evaluative Control: Agentic Activity without the Paradigm Features of Agency

In her paper, 'Two Kinds of Agency', Hieronymi (2009) is interested in delineating two different kinds of mental agency. The first kind can be understood as intentional agency and is linked to the type of control and agency present within intentional action. She highlights two features present within intentional action

⁸⁰Different conceptions of the agentic nature of judgments have been put forward by McHugh (2011) and Toribio (2011); both consider judgments to be a form of non-voluntary action that is, an action that does not have any intentional element. These types of accounts rely on different aspects of judgement to justify the claim that judgements should be described as actions. For Toribio (2011) judgements should be seen as actions because individuals are often held morally responsible for their judgements. She contests that since they form part of what an individual can be morally responsible for they should be considered a type of action. Under such a view it is the social role that judgements play that justifies their inclusion as actions. McHugh (2011) on the other hand suggests that judgements are actions because they are teleological, in that they aim at truth, however this aspect of judgments itself prevent them being voluntary because by aiming at truth they are constrained by it. I will not argue against these views, as understanding judgements as non-voluntary actions that are nonetheless agentic is not necessarily in conflict with seeing them as occurring under evaluative control, that is being moved by particular reasons, a view I will discuss forthwith.

that have been seen as essential for a mental episode or event to be seen as under our control. Firstly, an action has to be voluntary. According to Hieronymi, in order for an action to be voluntary, we have to be able to be moved to action, or to maintain an action, for any reason we take to count sufficiently in favour of carrying it out. Secondly, connected to this kind of voluntariness is the ability to stand at a *reflective distance* from the object of our intentions. To stand at a reflective distance is just, in forming an intention about what to do or in deciding to act in a certain way; we are required to take up a certain stance or mental position on the objects we wish to act upon. Hieronymi considers that the content of our thoughts, anything that is represented in an intention to act, as mental objects. When manipulating such objects, just like when we manipulate objects in the world, by acting intentionally we place a mental distance between ourselves and those objects. Those mental objects are viewed as distinct from us as the manipulator of those objects. It is this standing in relation to an object that allows us to act upon it. In taking up this reflective distance you are able to decide how to proceed with your action, for you have placed the object (mental or real-world object) in relation to you as an actor. These two features, voluntariness and mental distance, define for Hieronymi the basis for the first kind of mental agency that is consistent with the traditional view of intentional action (Hieronymi, 2009).

But, Hieronymi contends, there is a second kind of mental agency that does not depend on these features. Instead it finds its basis in a different notion of control: evaluative control. In order to motivate this second kind of agency Hieronymi (2009) utilizes a common enough conception of passive and active. Events that we are active in regard to can be held in contrast to events that we are merely passive in regard to, such as the experiencing of pain, or as events that happen to us, or that act upon us, like being punched or operated upon⁸¹. In making this distinction between being active and being passively acted upon. Hieronymi (2009) is able to propose a description of mental episodes that do not display, as she calls them, the paradigm features of agency, but yet are not passive. She suggests that we should consider this second kind of episode that includes the formation of beliefs and intentions as active, even though they do not possess the paradigm features of agency typical of intentional action (Hieronymi 2009).

Hieronymi utilises the term evaluative control to describe the personal activity of belief formation that is at the core of this second kind of mental agency. Evaluative control is the ability to reform or revise your take on an attitude. In the event of reforming our attitude to a question, we are exercising evaluative control if in revising that attitude it is altered in a way that is sensitive to our reasons. If our reasons for deciding were different, the revision of our attitudes would be appropriately sensitive to them (Hieronymi, 2009). In other words, exhibiting evaluative control is the ability to alter or revise our answer to the relevant question in a way that is controlled by or sensitive to our reasons for revising.

Evaluative control is not voluntary in the sense necessary for inclusion within the first tier of agency because as we have discussed above (subsection, 5.1.1) we are not free in deliberating to be moved

⁸¹ For more on passivity and activity, see Wittgenstein (1953), Anscombe (1957), Sartre (1958), Moran (2001), Mele (2009).

by any reasons, but only by reasons that are directed at truth. Further, in evaluative control there is no mental distance that separates us from the activity of our reasons (we will discuss this point momentarily in connection with Moran's (2001) concept of the deliberative stance).

In proposing this second tier of agentic activity, Hieronymi hopes to provide an area of agentic behavior that captures a type of mental activity that is bound up in our ability as reactors to respond accordingly to the phenomena we encounter, specifically to be agents that respond to reasons. In so doing she is not alone. Moran (2010), Smith (2005), Fischer and Ravizza (1998) all suggest accounts along similar lines. To return to the crucial point, these accounts of agency are not based on accounts of intentional action but are nevertheless accounts of agency, based on the control our reasons can have over our activities. Here, being responsive to reasons is being suggested as the core feature of agency. It is this feature that makes it possible for an individual to be held morally responsible, being an appropriate subject of praise or blame, for the outcome of those reason sensitive processes and they are appropriate candidates because controlling behavior in a reasoned way is the fundamental way individuals can be understood to exhibit agency. The mental activity associated with being sensitive to reasons is central to my claim that IBPS is active; we will spend some time getting to grips with it. To establish this claim I will turn to the philosopher Richard Moran (2001). His account, although complicated, is particularly rich and is perhaps the best way of understanding the agency involved in Hieronymi's account. In examining Moran's account it is worth bearing in mind that the basis of his claim that deliberating is agentic is based on an appropriate sensitivity to reasons.

5.3 The Deliberative Stance

Hieronymi's (2009) concept of evaluative control finds its inspiration in the work of Richard Moran (2001). Moran, in his book 'Authority and Estrangement' is interested in the consequences of adopting the first person perspective. He suggests that what is unique about the first person perspective is as much to do with agency and commitment as it is with any notion of infallibility, that our coming to decide on a particular course of action, or a particular belief, can be an agentic activity. He suggests that when we carry out our deliberation in a particular first personal way we are responsible for our deciding. When we deliberate on a matter and come to decide on what our answer to a particular question shall be, we are actively participating in that deliberation and can be responsible for it because in so deciding we change our attitudes and commit ourselves to a decision that is in accordance with our reasons. This first person perspective Moran calls the deliberative stance (the nature of which will be detailed shortly). Here, like evaluative control, the extent of our participation in the deliberation is limited to the change in the relevant attitudes brought about in accordance with our reasons for so deciding, and the commitments entailed by endorsing these new attitudes. We need to understand the nature of the deliberative stance for later I will make the claim that we are engaged in this deliberative activity within IBPS.

The deliberative stance is involved with a particular understanding of how we can have self

knowledge, and the basis of Moran's claim of how we can know what is going on in our own minds is the ability to avow our own state of mind. For Moran

An avowal is a statement of one's belief that obeys the transparency condition. (Moran, 2001 p.101)

and

A statement of one's beliefs about X is said to obey the Transparency Condition when the statement is made by consideration of the facts about X itself and not by either an "inward glance" or by observations about one's own behavior. (Moran, 2001 p.101)

It is worth our time here to explore Moran's grounding for this particular conception of transparency. Moran gives us this quote from Evans,

[I]n making a self ascription of belief, ones eyes are, so to speak, or occasionally literally, directed outward upon the world. If someone asks me "Do you think there is going to be a third world war?" I must attend, in answering him to precisely the same outward phenomena as I would attend to if I were answering the question "Will there be a third world war?" (Evans, 1982 p.225)

For Moran we are not to think that transparency here is the claim that the answer to the question 'whether I think there is going to be a third world war' and the question, 'whether there will be a third world war' are reducible to each other. For he suggests that we can have different answers to these questions (Moran, 2001, p.64). Rather what transparency amounts to here is a claim about the kinds of reasons that should be seen to play a role in deciding the outcome of a particular question one way or another. For Moran these two questions, the first concerning what you the agent will come to believe through deliberation on the subject, and the second about what the states of affairs in the world actually are, utilise the same set of reasons⁸². As Moran puts it a second time

The claim then, is that a first person present-tense question about one's belief is answered by reference to (or consideration of) the same reasons that would justify an answer to the corresponding question about the world. (Moran 2001, p.61)

So the answer to any question that is addressed to the actual state of affairs in the world, such as 'is it raining?' is transparent to the answer to any question brought about by consideration of the personal question 'should I believe it is raining?'. So when we utilise this group of reasons directed at the world we take up the position of the deliberative stance. For an agent to adopt the deliberative stance they do not actually need to ask the deliberative question in explicit terms. Since the deliberative question utilizes the same reasons as a world-directed question any world-directed question will automatically make logical

⁸² For a different take on transparency see Velleman and Shah (2005) and [Steglich-Petersen](#) (2006).

room for the individual to answer the deliberative question ‘what shall I take to be my answer?’ or ‘what will my answer be?’. As Moran says,

That is not to say that one normally arrives at one’s beliefs through some explicit process of deliberation. Rather what is essential in all these cases is that there is logical room for such a question ... (Moran 2001, p.63)

So for Moran when an agent asks a question that is addressed to the actual state of affairs in the world he is constrained to considering only reasons that are world-directed, but it is nevertheless transparent to a deliberative question about what the agent will take to be an answer to that same topic. It is the fact that both questions can only be answered by addressing the same reasons that makes the answer to both questions transparent to the other. Transparency makes it the case that whenever I answer a world-directed question I leave open the logical space to answer a deliberative question on the same topic. This for Moran is enough to consider the agent to be taking up the deliberative stance.

5.3.1 Why is the Deliberative Stance Agentive?

Within this thesis we have relied on the common assertion that for an activity to be agentive depends on it being controlled or responsive to the individual’s reasons. For Moran it is deliberating and coming to decide from the position of the deliberative stance that is the central way humans exhibit control and sensitivity to their reasons in constituting their mental states. In exhibiting that, reasoned control individuals can be seen to be morally responsible for their attitudes.⁸³ From the deliberative stance an individual is exhibiting control of his reasons because it is within deliberation that the person avows a particular attitude and through avowing comes to know his own mind on the matter. Individuals decide in the moment what they are to believe about any particular matter. Importantly that deciding is based on their reasons for what they should believe. As Moran says about the deliberative stance,

[I] address myself to the question of my state of mind in a deliberative spirit, deciding and declaring myself on the matter. (Moran, 2001 p.63)

It is important to note that the sense of activity that is gained within the deliberative stance is a long way from what might be termed voluntary control (as we discussed above, (subsection 5.1.1) understood as the ability to dictate the course of your deliberation for any reason you take to be in favour of so doing. When Moran talks about deciding in the moment, what is meant is that it is at that moment that a decision on the matter is made, not that you can decide what you want to believe for whatever reason you might want. Indeed, transparency suggests that you are restricted to the outcome of a deliberation that is identified with

⁸³ Moran prefers the term personal responsibility but I think it amounts to much the same thing. Both terms indicate the agent’s appropriateness for the ascription of praise and blame.

the specific group of reasons that are your reasons for deciding what you believe in accordance with the world. Again Moran,

‘to speak of the person’s role in forming his attitudes is not to invoke a kind of willful or wishful capacity for self creation. A person adopts this role insofar as he can answer questions of the sort ‘What am I to believe here?’ and thereby come to believe something, or answer a question of the form ‘is this what I really want?’ in terms of consideration of what is worth wanting and thereby come to clarify the structure of his actual desires. (Moran, 2001 pp.63-64)

For Moran we should consider the activity of the deliberative stance as specifically personal. For in answering the question in the moment ‘what shall I do?’, or ‘what shall I believe?’, there is a first personal reflection upon and commitment to a position about the world that can be said to be no-one’s but the deliberator’s and is formed from no reasons but his own. It is formed from a first person perspective that can only be occupied by the deliberator and is informed only by reasons provided by the matter at hand. It is on this active engagement of particular deliberative reasons that Moran (2001) bases his claim of agency. We can rightly talk of moral responsibility here because the deliberative stance shows how the process of deliberation is owned by the person doing the deliberating and that personal deliberation brings about a change that is responsive to the agent’s reasons for so deciding. This personal reason responsive activity then, that Moran outlines as the personal activity within the deliberative stance, is an important way of separating activity that belongs to an agent and is responsive to his reasons from a behavior that is not. An agent can be morally responsible for any belief formed in this deliberative way. Here Moran is resting his claim of agency on a bringing about change in an individual’s mentality that corresponds with his reason for deciding what he should, or shall, do or believe.

5.3.2 The Passive Nature of the Theoretical Stance Illuminates the Deliberative Stance

We can shed more light on the active nature of the deliberative stance outlined above by considering a different conceptual stance, the *theoretical stance*, that Moran considers can be applied by individuals to their own and others’ attitudes and beliefs. The point of outlining the theoretical stance here, as is Moran’s point in authority and estrangement, is to illuminate the active nature of the deliberative stance.

For Moran (2001) the theoretical stance is the taking up of a third person perspective about your own beliefs. It is the considering of your mental states as a third person observer. It is asking the question ‘what belief do I currently hold?’ not the deliberative question ‘what belief should I now adopt?’. You ask the former theoretical question from the third personal theoretical stance in the same way you can ask the question about another individual’s beliefs (‘Is it true that Tony believed we should stand shoulder to shoulder with the USA?’) and in examining your own beliefs from the theoretical stance you ask the same type of third person question directed at your own state of mind, at your own beliefs ‘Is it true that I currently hold the belief that we should stand shoulder to shoulder with the USA?’. Moran suggests that in

order to achieve this objective reflection it is essential to step back⁸⁴ or place a mental distance between the theoretical observer and your actual beliefs. A deliberation from the theoretical stance is based on the evidence assessable from the third person perspective and terminates in answering that question. So if you consider whether you are scared of snakes, you may take your inner states (such as your beliefs) or your behavior as evidence for such a judgement. In this sense you are making a judgement in a manner not dissimilar to the deliberative stance. But there is an important difference between this type of theoretical judgement and a deliberative judgement. A theoretical judgement never changes an attitude that you hold. You always come to a conclusion about how a current state of affairs stands from the perspective of an onlooker. That is, you examine a state of affairs to ascertain how it is now, not to decide on how it should be. In making such a judgement it cannot be the case that you change the nature of that belief or attitude, rather the theoretical deliberation is always addressed to assess the current character of a mental state. There is no capacity here for making a change in the state under assessment because the aim of your assessment is to ascertain what the current state of affairs is that you possess. So we are in a position of passivity towards the mechanisms we are observing and are therefore not active in regard to them. This is a necessary consequence of adopting the theoretical stance and directing our enquiry at understanding what the nature of those attitudes currently is, not at deciding what they should be. Your investigation does not aim at change or endorsement but only at obtaining facts. It is therefore not possible to separate the mental states that are the target of your investigation from the adoption of the theoretical stance. It is, however, from the theoretical stance that manipulation of mental states through indirect intentional control is possible. We will turn to examining this type of control in a moment.

This passive nature of the theoretical stance is in contrast to deciding from the deliberative stance where it is the case that you can bring about a change in your current attitudes. In deciding in a deliberative way you bring about that your attitudes reflect your reasons for thinking that you should hold a particular attitude and as such you will change your attitude or constitute it anew depending on your reason for thinking that attitude is the correct one to adopt. As Moran says

In characterizing two sorts of questions one may direct towards one's state of mind, the term 'deliberative' is best seen at this point in contrast to 'theoretical,' the primary point being to mark the difference between the inquiry which terminates in a true description of my state, and one that terminates in the formation and endorsement of an attitude. (Moran, 2001 p,63)

In occupying a particular stance the question you are engaged with is outwardly directed in the case of the deliberative question and inwardly directed in the case of theoretical question. The reasons you utilize in deciding a deliberative question can be considered your reasons. The fact that the deliberation in question

⁸⁴ See Korsgaard (1996), Sartre (1958), for similar talk of mentally stepping back.

is sensitive to these reasons in constituting an attitude provides the grounding for a claim of moral responsibility for that attitude.

The deliberative stance, then, can be seen as the position in which we normally reason and judge what it is we shall do within, or believe about, the world. It is the position of everyday reasoning, the first person position of considering for yourself whether or not to donate to charity, or to walk home, or emigrate to New Zealand. In adopting the deliberative stance our considerations are confined to a specific set of reasons. To return to our example of asking yourself, from the deliberative stance, the question ‘do I believe there will be a third world war?’. Importantly for Moran in answering this question of yourself from the deliberative stance, your answer becomes transparent to another question, that is, the question of deciding one way or another in accordance with the world, ‘will there be a third world war?’. The answers to these questions are transparent to each other, since both answers are arrived at by consideration of the same reasons. In reasoning in this way it becomes the case that our judgements are sensitive to the balance of our reasons. Our attitudes stand in a position that if new judgements are made or reasons brought forward our attitudes may be altered by that same process of deliberation. In taking up the deliberative stance we are morally responsible for the attitudes formed by our judgements. This process, of judging from the deliberative stance, I take to be one and the same as evaluative control. We have now set out a new understanding of the agentive nature of judgements. In a moment I will argue that this type of deliberative activity is present within the process of IBPS, but before then I want to consider a particular objection about seeing this type of evaluative activity as agentive.

5.4 An Indirect Objection

In chapter 3 we outlined the range of mental events that can be described as mental actions. These were limited to a small number of events that were describable as shepherding and prompting actions, and a broader range of events that can be understood as trying actions (section 3.4 and subsection 3.2.1). The range of mental actions was limited to a small set of events partially because of the access puzzle (subsection 3.2.3) and the consequent need to be able to represent the intended outcome within the content of the intention. Events such as the retrieval of particular memories and the process of IBPS were not capturable by these mainstream views of action and we concluded therefore that understanding mental agency in terms of mental action was problematic.

There’s an objection we did not deal with in chapter 3 but now, with the concept of evaluative control in place, we can answer. The objection goes like this: it is true that some mental episodes cannot be direct actions for the reasons you have described (sections 3.2 and 3.4), but what has been at the centre of the suggestion that intentional action is important for mental agency is not the direct type of voluntary control you described earlier (section 5.1.1), but a form of indirect intentional control, where we bring about the content of our intentions, not by directly acting on the mental state we wish to change, but indirectly by changing the world or other mental states, and in so doing bringing about a change in the

mental state we wish to control so that it conforms to our original intention. For example, if I want to believe in god, but after deliberation I find that I don't, I might be successful in forming that belief if I take Pascal's (1656) advice and adopt the manner and habit of those that do believe in god, and in this indirect way bring about the desired change in my beliefs. The objection here is that such indirect yet intentional actions contain one of the paradigm features of agency, mental distance, and can account for much of the agency concerning our mental lives, and that therefore there is no need to embrace a second tier of agentic activity based on evaluative control.

To examine this claim we need to look at how indirect intentional control works. Hieronymi (2009) suggests that we can take steps designed to indirectly bring about that we believe x. In order to succeed we have to bring it about that we have answered positively the question of x, i.e. that we form a commitment to x. There are many ways in which this is possible, from directly manipulating our environment so as to match the belief state we want to believe (Feldman (2000) points out that if you want to believe the light is on, you can bring this about by switching on the light), to subtle internal self deception through prompting, and a range of intermediates that will have different consequences for the agentic nature of the activity that brings about the commitment to x. Hieronymi's (2009) account of *managerial control* takes in a range of these possible scenarios. A further category of *manipulative control* is included to try to separate some worrying cases we might consider non-agentic.

In managerial control we control our mental episodes (judgements, intentions) indirectly. We form intentions to manipulate objects in the world specifically for the reason of bringing about a change in our mental states (Hieronymi, 2009). If you put up a sign on your office door about the advantages to your health and the environment of walking home rather than getting the bus and you intend when putting up the sign that it will change your attitude at the end of a hard day from intending to get the bus to intending to walk home, and it does in fact enter into a deliberation that has the effect of changing your attitude to match the attitude you intended to bring about, then you can be said to have exhibited managerial control over your attitude as to how you should get home. This is a form of indirect intention control because you have changed the world so as to bring about your desired intention, even if you have done so indirectly by altering the environment, not by directly changing the mental state you wish to alter. Managerial control then, is one way in which we can alter our beliefs in an intentional way.

There is an important proviso as to how useful this form of control is; even although managerial control is possible (in fact it is a technique commonly advocated in cognitive behavioural therapy⁸⁵) in everyday life the opportunities to managerially control your beliefs may be limited. This is the case because any reason you find compelling to alter your mind or behaviour has probably already formed part of your deliberative process and it is therefore not easy to provide yourself with extra reasons to change your attitudes (Hieronymi, 2009). Perhaps the case where managerial control may be most useful is

⁸⁵ See 'A Clinician's Guide to Mind Over Mood' Padesky and Dennis, 1995.

situations where two desires are in competition. For example if you like to smoke but you want to quit for health reasons, managerial control may be helpful in changing the balance of those desires by bringing to mind exactly why you support one over the other. In this indirect way managerial control can bring about changes in your judgements. You can be said to have controlled your judgment by forming an intention and putting in place a causal factor that is important in bringing about that change (Hieronymi, 2009). If we follow Moran (2001), it is also the case that if managerial control avoids the taking up of the deliberative stance managerial control is an indirect way of bringing about a result that matches your intention. The activity we are associated with is not found in actively deliberating rather we have set in motion events that act as an outside force upon our deliberation. When it is the case that we are active in regard to this event it is not because we have been active within the deliberation, but rather because we have manipulated ourselves into performing a result that matches our original intention. We have taken up a third person perspective that is similar in character to the theoretical stance (discussed above subsection, 5.3) and in so doing we have placed a distance between ourselves and our attitudes that allows us to be direct manipulators of those attitudes (Hieronymi, 2009). The theoretical stance is not the same as managerial control but managerial control is dependent upon it, as it is from the theoretical stance that such managerial ‘tinkering under the hood’ is possible. Importantly, the tinkering requires a separation that allows us to consider the mind as an object available for manipulation. Since we are in the position of viewing our attitudes and judgements as objects separate from us we can take up a position from which the inner workings of the mind can be examined. This mental distancing is vital for managerial control, since what we want to do is form an intention about specific mental states, we need to be able to intend to manipulate them.

Indirect intention control (one example of which is managerial control), then, can be seen to require mental distancing in the same way as it is required for direct (by which I mean not done by doing something else) actions. In order to do either we require to set an objective distance between ourselves and our mental states and this distance-making or stepping back is similar to the operation involved in taking up the theoretical stance. If we accept indirect intentional control in the form of managerial control, then, it seems we may have no need for evaluative control. So since indirect intentional control possesses one of the paradigm features of agency then we should abandon evaluative control. But here Hieronymi (2009) has an interesting line of argument. She contends that what is important about managerial control breaks down into two episodes of evaluative control.

5.4.1 Managerial/Manipulative Control as Two Instances of Evaluative Control

Within managerial control Hieronymi (2009) identifies two areas that she considers to be exercises of evaluative control. According to Hieronymi any successful exercise of managerial or manipulative control will require a commitment to an answer to two questions at two distinct points. Firstly, a commitment to a positive answer to the question of whether to manage or manipulate one’s attitude and secondly, a

commitment to whatever answer is embodied in the attitude successfully managed or manipulated. It is important to note two things. Firstly, as Hieronymi says, managerial control can be seen as an example of intentional control. That is, you can be said to be in control of it because you bring about a real world match with your intention and because the method of managerial/manipulative control (e.g. the sign on the door) is a causally relevant factor in your acting one way or another.

‘an exercise in managerial or manipulative control must be intentional to qualify as control, despite the fact that an exercise of evaluative control need not be. Managerial or manipulative control is a matter of acting so as to affect something according to one’s purposes. If one acts so as to affect something according to one’s purposes without intending to, it seems wrong to say that one has exercised control over that thing.’ (Hieronymi, 2009 p153)

Secondly, an important distinction should be made here on the basis of how the managerial control affects you. In the above explanation managerial control works by providing a reason that enters into your internal deliberation. But take another example, say, hypnotherapy. Here it seems that it is possible to intentionally control an action by it being imposed by such therapy but in this type of case the deliberation involved in evaluative control may overwhelm the previous stimuli (e.g. the hypnotherapy). Here Hieronymi suggests we should consider the title manipulative control more appropriate. She suggests that the boundary between the two might be blurry. ‘I suspect the distinction between management and manipulation will be hard to draw sharply’ (Hieronymi, 2009 p.153). It is right to say that in both cases we are required to deliberate about what we want to want; the commitment to this belief will be the same in both cases. It will also be true that in the event of successful manipulation/management once an outcome is forthcoming we will be committed to it, but it is not straightforwardly the case that this will be a matter of evaluative control. It will be a matter of evaluative control if this decision forms part of a further deliberation about whether you should or should not do something. Consider the door sign again. It can be correctly said, if on seeing the sign you enter a period of deliberation about what to do, that you have exercised managerial control. What is important is that the causal impetus of seeing the sign enters into your deliberation in a way that facilitates an outcome. Any outcome will suffice for that deliberation to be a product of evaluative control but for it to be an instance of intentional control it must be the case that the outcome matches the intention, so managerial control requires two exercises of evaluative control and a match between the intention and the outcome. But indirect intentional control may only require one instance of evaluative control and a match between the intention and the final outcome. Nevertheless, Hieronymi is right to state that intentional control can in some cases contain two instances of evaluative control but it is wrong to say that it must.

This is made clear if we look at a case of strong manipulation like the hypnotherapy case. For example, you want to start recycling but in the past have found it difficult to be motivated. A psychologist

friend of yours says he has a tape which if you listen to at night will make you recycle automatically without thinking about it. No second episode of evaluative control will be necessary as you will not have to commit to the outcome of the second deliberation because no deliberation will occur. You go to bed and set the tape because you intend to recycle the next day, and the next day you do recycle without considering the matter. I think this can be suggested as a possible candidate for indirect intentional action, and of manipulative control that does not involve two instances of evaluative control. So it seems it may be possible that an action can be intentionally controlled without having a second instance of evaluative control. What is essential for intentional control is the first instance of evaluative control. The forming of the original intention can be clearly seen as being dependent on evaluative control and necessary for intentional control. It is for this reason that I agree with Hieronymi when she says that intentional control cannot be clearly separated from evaluative control, and if we follow this line we should not consider intentional control as superior to evaluative control for it depends upon it on at least one, if not two key junctures. The indirect objection, then, loses its force, for even if indirect intentional control contains one of the paradigm features of agency it relies on evaluative control on at least one occasion and so evaluative control should be considered an important form of agency⁸⁶.

5.5 The Story so Far

So far I have shown that we are only able to control our mental activity through intentional action construed according to the simple view (subsection, 3.2.6) or GCT (sections, 3.4) in a small number of cases. One restriction on the intentional control of mental episodes is the requirement that your intention contains, up front, content that represents what you wish to bring about. But there is a further restriction in the case of judgments in that it is not possible, even if we have the content upfront, to intentionally judge *x*, where *x* is a specific outcome (subsection 5.1.1). To make a judgment about what we believe is the case requires that the specific outcome of that judgement, belief in *x*, results from the balance of our reasons for

⁸⁶ We might think that managerial control could provide an account of agency, for the process of IBPS. To pursue such an account some accommodations similar to those suggested in chapter 3 would have to be made (section, 3.2.2 and 3.4). Since managerial control relies on intentional control we will have to propose that an intention to problem solve, or to think about a problem at a later date, occurs within preparation. The need for an intention returns us to the access puzzle and the difficulties we faced in providing an account of intentional action for IBPS (section, 3.2.2 and 3.4). Without knowing what the outcome of IBPS is, it will not be possible to intend to bring it about. It will not therefore be possible to managerially control the content of the problem solution because we are unable to represent that content within our intention. At best managerially controlling IBPS will provide an account of indirect action for the bringing about of coarse grained intentions e.g. 'trying to solve a problem'. Such an account of IBPS is subject to the same worries as accounts based on intentional action, that it is not a natural description of the process, that it is over-intellectualized and that it will only apply to some cases (section 3.4). Although it may be possible to provide accounts of agency for IBPS based on managerial control any such account would seem to suffer from the same disadvantages as accounts of agency for IBPS based on intentional action. Most importantly, we will be unable to capture the bringing about of the content of particular problem solutions as under the influence of managerial control.

so judging.

We have also discussed the possibility of intentionally controlling our mental episodes indirectly through a process of managerial control. As we discussed above (section 5.4.1) such an approach will be limited in several ways. Managerial control is limited to circumstances where we can provide additional reasons for coming to a judgement. Further, the difficulty in successfully completing an episode of managerial control limits its application. There is, after all, a finite number of ways and situations where we have the time and space to execute such control.

On the other hand, evaluative control provides a way of gaining responsibility over mental episodes through the occupation of the deliberative stance (sections, 5.2 and 5.3). Next I will examine whether evaluative control, or as I will say occupation of the deliberative stance, can provide a useful account of agency for the process of IBPS.

Moran's argument for understanding the personal nature of the activity associated with deliberation has been set out in full (sections, 5.3, and 5.3.1). Further I have argued that deliberations of this kind cannot be seen as intentional (section, 5.1.1) Below, some of the difficulties associated with this account are discussed in terms of the deliberative stance being applicable to IBPS but no general attempt to defend Moran's position in the larger philosophical debate has been carried out. Rather, I will adopt Moran's view as a coherent, mainstream position that we have reason to think is vindicated by the new understanding of IBPS that adopting it provides. If Moran's account of the deliberative stance is correct, IBPS should be seen as involving the deliberative stance. This view will provide an account of why IBPS is active and agentic. In turn this account of IBPS opens a new window on the type of deliberative activity Moran has in mind. I will proceed, then, by arguing that within IBPS the stage of revelation - the moment of insight itself - is a deliberative judgment.

5.6 Deliberative Activity within IBPS

In chapter 2 (section 2.4.4) we concluded that the process of IBPS includes four stages. Firstly preparation, where different processes like problem finding and information gathering are unitized but are unsuccessful in delivering an answer (subsection 2.4.1). This is followed by incubation, a period without intentional activity that is functionally important in bringing about revelation (section 2.4.2). Revelation is the stage where the new insight comes to mind (section 2.4.3). It is followed by verification, where the insight to the problem at hand is integrated and an answer produced (2.4.4). Earlier in this chapter I identified evaluative control as the central form of control associated with deliberation and the making of judgements (sections, 5.2 and 5.3). I will argue that evaluative control can provide a useful way of understanding the activity of IBPS. In the rest of this chapter I will focus on providing a detailed account of the activity of IBPS through the occupation of the deliberative stance. In providing this account I will have accomplished one of the main aims of the thesis, for the nature of the agentic activity associated with IBPS will have been established.

5.6.1 Evaluative Control and IBPS

Within the first part of this chapter we outlined the nature of the activity associated with the first person perspective, particularly within the process of deliberation. In thinking along these lines we were in broad agreement with Moran (2001), Hieronymi (2008, 2009) and Smith (2005). But that is not to say that such a view is universally accepted (section, 5.1.1). In the rest of this chapter I will accept this line of thinking and suggest that some of the important stages of IBPS can be understood as active in this deliberative way. This will not only provide an account of the agential nature of IBPS but also highlight the nature of the non-intentional deliberative activity that Moran has in mind. I will begin by examining whether the underpinning concept within evaluative control, Moran's concept of the deliberative stance, can be occupied by the individual within the process of IBPS. This will involve identifying areas within IBPS where the key features of the deliberative stance are present.

As we set out above (section 5.3), occupying the deliberative stance necessarily involves the agent, for the agent utilizes a particular set of reasons for deciding and avowing one way or another on a particular subject. The avowal of the agent, his coming to decide in accordance with his reasons for deciding, is an exhibition of agency. It is the personal activity of the agent. The process of rational deliberation can be seen as the personal activity of the agent in a way in which the activity of other mechanisms are not. It is deliberating in the active manner inherent in the position of the deliberative stance i.e. in accordance with the reason for thinking what shall I believe, that will be central to seeing the process of IBPS as active, as occupying the deliberative stance and therefore acquiring the authority and control that the deliberative stance affords.

5.6.2 The Deliberative Stance in IBPS

I will now argue that within IBPS you are taking up the deliberative stance, and are therefore active at two key stages, preparation and revelation. In preparation, you are involved in a number of activities such as problem finding, information gathering, and concept selection (section 2.4.1). Several of these processes may involve deliberation. You are also deliberating, trying to find a solution to your problem. For the moment I will be concerned with the straightforward deliberation about the question to which the individual is seeking an answer. This deliberation is particular to preparation. We are trying to solve a particular problem but the solution eludes us. Here, the individual engaged in IBPS is asking a world-directed question; what is the answer to the particular question at hand, e.g. why does light bend? The reasons that are in play in your consideration of the question are just those that are relevant to answering the question at hand. At this point, you are addressing the question 'what is the answer to this question?', the answer to this question is transparent to the answer to the question 'why should I believe light bends?'. You can therefore properly be described as occupying the deliberative stance and being active in regard to that deliberation.

Here, however, the similarity between belief formation and the deliberation during preparation breaks down. For in the case of preparation what has the agent been active in producing? In the case of belief formation the end result of the deliberation that we are active in producing is the avowal that this is what we shall believe. But in preparation there is no obvious immediate positive outcome. Neither do we necessarily come to a conclusion of our deliberation. However this might not be a difficulty for as Moran (2001) suggests, it is not the case that we need to conclude our deliberation for us to be active in regard to that deliberation, for we make a commitment to our deliberation as it progresses. So wherever our deliberation stops we gain a conviction about the outcome of that deliberation. What is important is that you gain conviction in the form of an intention or belief. But it does not seem that this will necessarily be a conviction that a particular outcome is beneficial or useful or correct, in as much that you will have to gain a way to move forward, but only that you are convinced by the position of your own reasoning. If we agree with Moran, interruption in our deliberation may not be problematic for considering those deliberations active.

5.6.3 Objection 1: Are the Processes Involved in Preparation Intentional?

As we examined in chapter 2, preparation can be seen as a mixture of different processes. As indicated by the experiments described in that chapter, subjects can be asked to intentionally engage in problem construction, information finding, and concept selection. It seems that in everyday life, individuals can also mentally prompt themselves to intentionally engage in these processes. One legitimate concern we can have is that the activity of preparation should be seen as a form of intentional engagement or action. It is clear, however, that individuals involved in IBPS must also be involved in actively deliberating within preparation. The fact that intentional agency may also play a role does not detract from the deliberative activity that is taking place. Further if we agree with Hieronymi (2009), any instance of intentional activity also includes episodes of evaluative control that can be understood as the type of deliberative activity Moran has in mind. I think for these reasons we should consider that within preparation the agent is occupying the deliberative stance and can be afforded the sense of activity that the stance provides.

5.6.4 Incubation and the Activity of the Deliberative Stance

What about the rest of the process? The low level reportability of incubation makes it a poor candidate for being seen as active. However, on Moran's (2001) account this is not necessarily a problem as both conscious and non-conscious processes could theoretically be instances of the deliberative stance, even if in practice the high level nature of the processing normally provides reportability. What is essential in Moran's view is the distinction that the activity within the deliberative stance is the interplay of your reasons that leads to your altering or committing to a belief in virtue of those reasons for deciding what you shall in that moment come to believe. It is the activity of your reasons in so deciding that for Moran (2001) can be seen as personal. This is not dependent on a notion of conscious accessibility but rather on

the bracketing of a particular grouping of reasons, your reasons for your answering of the question of what you shall believe, as identifiable with the personal level activity of the agent (2001). What is essential, then, for Moran is that in occupying the deliberative stance we deliberate from a first person perspective and come to change our attitudes based on our reasons for so deciding. If we are unable to ascribe reasons to the activity within incubation then we are not in a position to ascribe to it the activity afforded by the deliberative stance. There seems no reason why incubation should be considered active in this way.

5.6.5 The Deliberative Nature of Revelation

The conclusion of incubation is the production of revelation: the third stage of IBPS where new content is brought to mind accompanied by the Aha experience. Embedded within the descriptor revelation is the event of revealing, the sudden uncovering of a product. This gives the impression that revelation is entirely passive, that something is revealed to us. Consequently, we may feel that revelation is entirely passive, that it is just what happens to us as a result of preparation and incubation. I want to persuade you that this is incorrect. To do so I will interpret revelation in terms of the active / passive distinction revealed by the deliberative stance. That revelation itself, exactly that moment of the eureka, can be seen as active, for if we understand it correctly we should not only see it as the uncovering of unconscious processing to the conscious mind but also as a judgment, a deliberative judgement that the product of incubation is my answer to the question laboured over in preparation.

Let us then take a closer look at revelation. Here we encounter the unique phenomenology of the creative process - 'out the blue it came to me, that eureka moment'. In a way, what could be more alien; to what could we be more passive in an intentional sense? But there is another side to the phenomenology; together with this feeling of being struck is the feeling of realization that what has struck me is an answer to the question I have been toiling over. It is this aspect of revelation that I hope to persuade you is active in the deliberative sense. The groundwork for this has already been put in place because in taking up Moran's (2001) distinction of passive and active we can see the event of revelation as involving occupying the deliberative stance. As stated above we are occupying the deliberative stance when we are deliberating on the question 'what shall I believe?', where as Moran says, a statement obeys transparency when we only refer to facts about the world and not facts about our own behavior (section, 3.1). To see this, note that along with the feeling of surprise and elation, there is another side to the phenomenology of revelation, which, as noted by Topolinski and Reber (2010) (section, 2.2), is a conviction on behalf of the problem solver about the truth of the product revealed to her. It seems that the confidence we gain is directed at the specific content of the revelation in relation to it answering the particular question we have been deliberating over in preparation. This realization of the product of incubation can be seen as a judgement. It is an affirmative judgement that the product produced in incubation is an answer to the question deliberated over in preparation. This aspect, I claim, is active in the deliberative sense.

We need to examine whether the realizing moment or judgement, that the product of incubation

answers the question the problem solver is considering within preparation, fits the criteria of the deliberative stance. What is required for this to be the case is that the judgement in question is a deliberation that is determined by the reasons specific to the consideration of what you, the deliberator, shall believe (Moran, 2001). The answer to this question is transparent to the answer to the world-directed question 'is this the case?', for it is decided with respect to the same set of reasons. In virtue of being decided by the same set of reasons, answering the question 'is this the case?' leaves logical room for the deliberative question 'what shall I take to be the case?', as such you are occupying the deliberative stance whenever you utilize this set of reasons.

In the case of revelation, the question at hand is whether the revealed product answers the question you were trying to solve in preparation. For this deliberation to be within the deliberative stance, the deliberation must be restricted to the reasons that are appropriate for answering the question of whether this is indeed an answer to the question you were deliberating over in preparation. This must be the case. Your deciding, in the moment of revelation, that the product is an answer to the question at hand must require you to consider whether the particular content is an answer to the question at hand. It, then, is arrived at only through considerations of facts about whether this is indeed an answer to your question. Thus your coming to the conclusion that this particular content *is* an answer is not under your ability to voluntarily choose one way or the other. Your deciding here is restricted by your reasons for so deciding in accordance with the world. In the making of this judgement the agent is occupying the deliberative stance because of the transparent nature of the answer to this question to the answer of the deliberative question, 'what shall I take as an answer to my question?', and so in utilizing the same set of reasons it leaves open the logical space to answer the deliberative question.

As we have discussed, for Moran (2001), answering a question in this deliberative way is a personal activity. The activity of such deliberation should be considered the activity of the agent, an activity that he can be held morally responsible for. The particular activity that can be described as personal in this case is the production of the IBPS product, for it is in judging the product of incubation, as an answer to the question at hand, that it becomes available to us for further scrutiny and action. Before that point, before that realization and judgment, the product is just another piece of unsubscribed content.

5.6.6 Objection 2: Is There Really a Judgement within Revelation?

One preliminary objection is whether there is a judgment going on within revelation at all. Is it possible to judge whether or not the product of incubation is an answer to the question under consideration in preparation, without a straightforward prior criterion of what an answer would involve? In looking at preparation, the problem solver is in specific pursuit of a problem. So we have some criteria in mind for what a solution may consist of. Some of these may be explicit and some may be implicit. The detailed nature of what is or is not a solution may very well not be present in your mind. You may have only a partial grasp on the nature of your problem. Yet this should not be considered a difficulty for judging

whether or not you have found a solution. For whatever notion you have of the problem will be sufficient for making the assessment that the new content answers the question at hand with the solution, and what can this assessment be but a judgement? What we have is an interplay of reasons appropriate for answering the question 'shall I take this as answering the question?', and if we have that, then we cannot but be within the deliberative stance.

This activity of revelation is possibly unique, for it is an example of deliberative activity that has absolutely no intentional element. Here I am not intending to decide, rather I am actively deciding in accordance with my reasons for deciding. These reasons are restricted to the reasons for believing that are in accordance with the world (subsection 5.1.1). This is unlike preparation and standard episodes of deliberation, where there is a mixture of intentional and deliberative activity. As we discussed in subsection 5.1.1 judgements may include shepherding and prompting actions that have intentional elements even if the making of a judgement is not an intentional action. On the other hand the activity of revelation presents a clear example of deliberative activity that cannot be interpreted as in any way intentional. This episode of active deliberation within revelation, in containing no intentional element, provides a clearer separation between the type of upfront activity of shepherding and priming action (section, 3.2.1) that Strawson has in mind and the deliberative activity Moran is concerned with, both of which we might consider occur within typical cases of deliberation. But in revelation we have a case of deliberative activity without any intentional involvement and as such it provides a way of examining the type of activity Moran has in mind without that examination being clouded by intentional elements.

5.6.7 Objection 3: Quick Judgements

A further objection to our position that revelation is a judgement is that it occurs too quickly to be deliberative. Shoemaker (2003) provides some examples of the type of quick judgements that might be problematic for Moran. Shoemaker's examples of quick judgements are judgements about the content of the recent past, such as 'Did you just come from your office?', or judgements where it seems the answer is immediately available, such as 'Are you wearing pants now?' We can express our concern (not specifically Shoemaker's concern) this way: since it seems that the answer to the question you are asking yourself is directly available to you, there is no need for deliberation; in this case it cannot be correctly said that you were occupying the deliberative stance in providing an answer, and so you are not personally active within the process.

The examples above may pick out judgements where we do not need to deliberate. Why is this the case? It seems that we have access to the answer in a direct way. In the example concerning the wearing of pants this may be the case because we are, on some level, always aware of their presence. Our sense of wearing clothes is adopted into our sense of bodily awareness. Consider an example of someone who has no sense of touch. In this case it might be that he has to deliberate, or at least check with the world through another modality to answer the question, 'Are you wearing pants?' - 'Well, I remember

putting them on this morning!'.

The applicability of direct awareness might be the same in cases where a memory is still so fresh in the mind as to make the answer obvious. But it is not clear when asked even an elementary question like 'What is your first name?' that you will be in a position to immediately respond. Of course, this does not rule out that sometimes you are in a position to immediately respond. It seems that information that I do not have immediate awareness of will require a judgement in retrieving. It is worth considering whether if you are already aware of some content and you base your answer on that content, then you are, in a way, not making a judgement at all, you are just accessing the information. But I think this is wrong. It will still be the case that you have to make a judgement about what part of the content that you are aware of is appropriate for your answer. Consider the question, 'Are you wearing pants?' You might think that the answer to this is immediate and does not require a judgement. It seems, however, that we have to distinguish between the different contents we have immediate access to, for example, that I am wearing shoes, that I am in my office, that I have a hat on, and so forth. So we may need to make a judgement that this is the correct bit of content that I have direct awareness of, that I draw on when I supply the answer, 'Yes, I am wearing pants'.

This brings us to a deeper point, that this type of quick judgement may not be a deliberation that connects to your reasons for deciding that you are wearing pants. This is not to say that you do not have reasons for wearing pants. There are many situations, almost all occasions when you wear pants, where you have reasons (even if those reasons are bad ones) for wearing pants; rather, it is to question whether when you come to answer the question 'am I wearing pants?' you utilise the set of reasons relevant to the deliberative stance, i.e. those reasons that form the evidence for you deciding whether or not you are wearing pants.

How could it be the case that this type of judgement connects with a deliberation? Most directly you could think that the judgement itself connects with reasons for thinking that the feeling you are wearing pants is the correct piece of content that you are aware of. That is, in answering the question and identifying the correct content, you are making a judgement that you are wearing pants for the reasons you believe you are wearing pants. The judgement necessarily involves taking the feeling of wearing pants to be a reason for forming the belief that you are wearing pants. This kind of quick judgement is therefore a deliberation, and does connect with a set of reasons that you believe you are wearing pants. The answer to this world directed question is transparent to the answer to the deliberative question on the same topic. So it leaves logical room for answering the deliberative question of whether you believe you are wearing pants. As with any other deliberation of this sort you cannot decide beforehand whether or not to give an affirmative answer to this question but rather it is dependent on the balance of your reasons in accordance with the world to directly answer the question that you are wearing pants.

In considering IBPS we can now see that the speed of the judgement involved in IBPS will not be a problem is describing the activity of insight as occupying the deliberative stance. Even if the judgement

is quick it provides the agent, if he is so asked, with a rationale for answering the deliberative question as to why he took that content as answering the question at hand. This logical space will always be available if the question he is answering, in judging, is a world-directed question, e.g. 'is this an answer to the question under consideration?'. For the answer to such a question is transparent to the answer to the deliberative question 'shall I take this as an answer to my question?', because they share the same reasons and therefore an answer to the world-directed question will necessarily leave room for an answer to the personal deliberative one. As such, the activity of revelation can be seen to occur through the occupation of the deliberative stance.

If we agree with Moran, it is now clear that we are active within the preparation and revelation stages of IBPS. We have provided strong justification for the intuition that IBPS is a matter of agency. Individuals engaged in IBPS are actively deliberating at two key junctures: up front in preparation, and in revelation. The range of mental processes that can be considered the activity of the agent, and for which they can be morally responsible, must now include the revelation stage of IBPS. We have established, then, one of the main goals of the thesis (section, 0.6): an account of the agency that is fundamental to particular aspect of creativity, insight based problem solving.

5.7 Bringing it all Back Home: Conclusions on the Agentive Nature of IBPS

Up to this point I have presented the case for understanding IBPS as containing two episodes of deliberation. I have argued that in IBPS individuals occupy the deliberative stance within preparation and revelation and therefore that these stages contain episodes of agentive activity. Having established these key points I want now to consider the importance and consequences of this position that I will call the deliberative account of IBPS.

The view that IBPS is a deliberative activity has implications for our understanding of the very nature of this type of deliberative activity. It opens a window on this deliberative activity as a type of personal agentive activity that unlike other episodes of deliberation is not obscured by intentional elements. Further understanding IBPS as agentive has important consequences for understanding creativity as agentive. I will explore these consequences in the conclusion. Before then I will examine the consequences for mental agency of understanding IBPS as agentive.

As I have argued, IBPS is agentive both in revelation and preparation. These two episodes can be seen as the personal activity of the IBPS agent. Out of these two episodes of deliberative activity the judgement in revelation is the most crucial for our understanding of the agency within IBPS. For in understanding revelation as active we can say that the activity of bringing new content to mind, to resolve the cognitive problem faced by the problem solver, is deliberative. It is us, the problem solver, who comes to judge the new content as relevant to the problem under consideration, and so the bringing about of new content, revelation, can be seen as the activity of the agent. The agent can be said to engage in actively constituting the contents of his mind in a way for which he can be morally responsible for. This provides

an explanation of why we are agentive for the specific products of IBPS. In chapter 3 and 4 we were unable to provide an account of agency for these products in terms of intentional action or consequences of intentional actions. In chapter 4 we discussed the importance of being able to make such an ascription (section 4.6). These products can now be seen to lie soundly within the scope of our agentive behavior and we can be considered to be morally responsible for the bringing about of these products. Individuals engaged in IBPS are therefore eligible for praise and blame for the content of the solution, i.e. whether it is a good, useful, creative answer to the question or not. In providing an account of the deliberative activity of revelation the picture Strawson presents, that mental agency is perhaps limited to up front intentional action, should now be expanded to include the deliberative way content is brought about. For we now have a clear sense of the deliberative activity that is taking place within a context that is devoid of any intentional activity.

When we discussed the minimal creative agency account (section, 4.5) we suggested that accounts of agency based on moral responsibility for action consequences may not be as satisfactory as accounts based on specific episodes of behavior that are controlled by or sensitive to reasons. We have now characterized the episodes of agentive behavior within IBPS as deliberative, or as we can say, under evaluative control. The agency involved in IBPS can be understood as being constituted by specific episodes of deliberative activity. This is distinct from the account of minimal creative agency we presented in chapter 4 (that finds its basis in Stokes (2011)), where actions in preparation can be considered as intentional actions and the further ascription of agency for the latter stages of IBPS is dependent on these intentional actions. As we discussed in chapter 4, such accounts will not provide a claim of responsibility for the bringing about of specific IBPS products, it is this specific content that will form any assessment of creativity in being judged novel and useful. This, for me, is the major deficiency in that type of account. It is a deficiency that is addressed by understanding that the moment of insight, the stage of revelation, is deliberative.

As I suggested above, being morally responsible for action's consequences is not dependent on those consequences being episodes of activity that are controlled by or sensitive to reason such as is the case with intentional action. Accounts of agency based on the direct results of intentional actions are more satisfactory than accounts of agency based on being morally responsible for action consequences because the former can be seen to bring about results in accordance with our reasons for acting while the latter cannot. The deliberative account of IBPS shares this feature with intentional action, in as much as both are sensitive to reasons. So the deliberative account seems on this basis to be more satisfactory than an account based on action consequences.

However one point that can be considered in favour of the minimal creative account is its foundation in intentional action. Accounts of agency based on intentional actions or the consequences of those actions have a traditionally strong place in accounts of agency. They share what has been termed the paradigm features of agency, that they are voluntary and require an agent to adopt a mental distance

between themselves and the object of their action (See Heironymi, 2009, section, 5). Accounts of this type fit better with a traditional picture of agency and we might therefore consider that they are in this way superior to the type of deliberative account I have established in this chapter. However we may feel that the fact that the deliberative account provides an account of agency for the products of IBPS to be an important difference between the two accounts. It is worth abandoning the paradigm features of agency to gain the new understanding that the deliberative account provides. Hieronymi (2009, subsection, 5.4.1) makes the point that the nature of intentional action is based on two episodes of evaluative control. If this is the case then we might consider that the importance of the paradigm features of mental agency has been overstated, and what is at the core of mental agency is evaluative control. If we follow this line of argument then the deliberative account, in taking this point on, provides an account of the agency that gets at the fundamental nature of the agency that is present in IBPS. The deliberative account then can be seen to provide an account of how specific IBPS products are brought about in an agentic way that captures the actual activity of the agent in a way that provides new understanding of the process of IBPS.

Understanding IBPS as agentic may also help our understanding of the agentic nature of deliberation. In IBPS the stage of incubation, a stage where no intentional activity is undertaken, is followed by the revelatory and deliberative active experience of insight. This is distinct from other types of deliberative activity where intentional activity is more closely associated with the deliberative activity and the distinction between the two types of activity is less clear cut. On the contrary, the deliberative activity of IBPS in revelation could not be more separated from any intentional activity, as such it demonstrates the distinct nature of the deliberative activity that Moran has in mind. This kind of deliberative activity can then be seen to be entirely devoid of the notion of intentional action, of forming an intention to change the world in a way desired by the actor. The activity of deliberation, which Moran assigns as the foundation of the personal, the foundation of what it means to be a free thinker, to constitute for ourselves our beliefs and character, has at heart no intentional element but rather is the activity of coming to decide on a matter in a way that is directed by the weight of the evidence presented by the world. Understanding the deliberative activity within revelation clarifies this point. It highlights the particular character of deliberative activity, devoid as it is of any intentional component, tuned rather to the coming to decide, in accordance with the world, what you shall believe. That is still, necessarily, the personal commitment of the agent and is central, for Moran, in forming the basis of the moral responsibility of the person. In identifying active deliberation within IBPS I have shown it to be within the range of processes that are fundamental to our conception of what it is to be human, what it is to be an agent, to be a being that can demand of itself a commitment to particular attitudes and beliefs that are both self affirming and constituting.

Bibliography

- Adams, F., 1986. Intention and Intentional Action: The Simple View. *Mind & Language*, Vol.1 (44), pp. 282-301.
- Amabile, T. M., 1982. The Social Psychology of Creativity: a Consensual Assessment Technique. *Journal of Personality and Social Psychology*, Vol. 43 (5), pp. 997-1013.
- Amabile, T. M., 1996. Creativity in context: Update to "The Social Psychology of Creativity". Boulder, CO: Westview Press.
- Amabile, T. M., Conti, R., Coon, H., Lazenby, J. and Herron, M., 1996. Assessing the work environment for creativity. *Academy of Management Journal*, Vol.39, pp. 1154-1184.
- Anderson, B. F., 1980. *The Complete Thinker*. Englewood Cliffs: N. J. Prentice Hall.
- Andreasen, N. C., 1987. Creativity and mental illness: Prevalence rates in writers and their first-degree relatives. *The American Journal of Psychiatry*, Vol. 144 (10), pp.1288-1292.
- Anscombe, G. E. M., 1957. *Intention*. Ithaca: Cornell University Press.
- Aristotle, 1984. *Poetics*, Chapter 17, 1455a pp. 32-34. In: *The Complete Works of Aristotle: The Revised Oxford Translation*, Vol. II, ed. Jonathan Barnes, Princeton University Press.
- Aristotle, 1993. *Nicomachian ethics*. Athens: Kactus Press.
- Ash, I. K. and Wiley, J., 2006. The nature of restructuring in insight: An individual-differences approach. *Psychonomic Bulletin & Review*, Vol. 13 (1), pp. 66-73.
- Audi, R., 1993. *Action, Intention and Reason*. Ithaca: Cornell University Press.
- Audi, R., 2001. Epistemic virtue and justified belief. In: A. Fairweather and L. Zagzebski, eds. (2001). *Virtue Epistemology: Essays in Epistemic Virtue and Responsibility*. Oxford University Press. pp. 82-97.
- Austin, J., 1863. *Lectures on Jurisprudence, or the Philosophy of Positive Law*. Ed. Robert Campbell. New York: James Cockcroft, 1875. First published 1863.
- Bailey, C. H., Bartsch, D. and Kandel, E. R., 1996. Towards a Molecular Definition of Long-term Memory Storage. *Proceedings of the National Academy of Science, USA*, Vol. 93, pp. 13445-13452.
- Bailin, S., 1984. Can there be creativity without creation? *Interchange*, Vol. 15, pp.13-22.
- Bargh, J. A., Chen, M. and Burrows, L., (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of personality and social psychology*, *Journal of Personality and Social Psychology*, Vol. 71 (2), pp. 230-244.
- Beeftink, F., Van Erde, W. and Rutte, C. G., 2008. The Effect of Interruptions and Breaks on Insight and Impasses: Do You Need a Break Right Now? *Creativity Research Journal*, Vol. 20 (4), pp. 358-364.
- Beghetto, R. A. and Kaufman, J. C., 2007. Toward a broader conception of creativity: A case for "mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts*, Vol. 1 (2), pp. 73-79.
- Bennett, J., 1990. Why Is Belief Involuntary? *Analysis*, Vol. 50 (2), pp. 87-107.
- Bentham, J., 1789. *An Introduction to the Principles of Morals and Legislation*. Ed. J. Burns and H. Hart. London: Methuen, 1970. First published 1789.
- Berlin, I., 1969. *Four Essays on Liberty*. New York: Oxford University Press.
- Boden, M. A., 1990. *The Creative Mind*. London: George Weidenfield and Nicolson Ltd.
- Boden, M. A., 1994. What is Creativity? In: M. A. Boden, ed. 1994, *Dimensions of Creativity*. Massachusetts: MIT Press. 75-118.
- Boden, M. A., 1995. Modelling creativity: reply to reviewers. *Artificial Intelligence*, Vol. 79, Issue 1, pp. 161-182.
- Boden, M. A., 2003. *The Creative Mind. Second edition*. Routledge.
- Boden, M. A., 2009. Computer Models of Creativity, *AI Magazine*, 2009, pp. 23-34.
- Bowden, E. M., Jung-Beeman, M., Fleck, J. K., Kounis, J., 2005. New approaches to demystifying insight. *Trends in Cognitive Sciences*, Vol. 9 (7), pp. 322-328.
- Bratman, M., 1984. Two Faces of Intention. *Philosophical Review*, Vol. 93 (3), pp. 375-405.
- Briskman, L., 1980. Creative Product and Creative Process in Science and Art. *Inquiry*, Vol. 23 (1), pp. 83-106.
- Cox, C. M., 1926. *The Early Mental Traits of 300 Geniuses*, California: Stanford University Press.
- Crockenberg, S. V., 1972. Creativity Tests: A Boon or a Boondoggle for Education. *American Educational Research*, Vol. 42, pp. 27-45.
- Csikszentmihalyi, M., 1999. Implications of a Systems Perspective for the Study of Creativity. In: R. J. Sternberg, ed. 1999. *The Handbook of Creativity*. New York: Cambridge University Press. pp.313-335.

- Danko, S., Starchenko, M., and Bechtereva, N., 2003. EEG local and spatial synchronization during a test on the insight strategy of solving creative verbal tasks. *Human Physiology*, Vol. 29, pp. 129-132.
- Davidson, D., 1963. Actions, reasons, and causes. *Journal of Philosophy*, Vol. 60(23), pp. 685-700.
- Davidson, D., 1969 The individuation of events. In N. Rescher (1969) *Essays in Honor of Carl G. Hempel*. Reidel.
- Davidson, D., 1980a. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Davidson, D., 1980b. Intending. In: D. Davidson (1980) *Essays on Actions and Events*. Oxford: Oxford University Press.
- Davidson, J. E., 2003. Insights about Insightful Problem Solving. *The Psychology of Problem Solving*, Davidson J. E., & Sternberg, R. J. (Eds.). (2003). *The psychology of problem solving*. New York: Cambridge University Press.
- Dennett, D., 1984. *Elbow Room: The Varieties of Free Will worth Wanting*. Oxford: Clarendon Press.
- Diertrich, A., and Kanso, R., 2010. A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin*, Vol. 136 (5), pp. 822-848.
- Dominowski, R. L., 1981. Comment on "An examination of the alleged role of 'fixation' in the solution of several 'insight' problems" by Weisberg and Alba. *Journal of Experimental Psychology: General*, Vol. 110 (2), pp.193-198.
- Dominowski, R. L. and Buyer, L. S., 1989. Retention of solutions: It is better to give than to receive? *The American Journal of Psychology*, Vol. 102 (3), pp. 353-363.
- Dominowski, R. L. and Buyer, L. S., 2000. Retention of problem solutions: The re-resolution effect. *The American Journal of Psychology*, Vol. 113 (2), pp. 249-247.
- Dorsch, F., 2009. Judging and the Scope of Mental Agency. In: *Mental Actions*, L. O'Brian and M. Soteriou, eds. 2009. New York: Oxford University Press. pp. 38-71.
- Elster, J., 2000. *Ulysses unbound: Studies in rationality, precommitment and constraints*. New York: Cambridge University Press.
- Engel, P., 2002. Volitionism and Voluntarism about Belief. *Croatian Journal of Philosophy*, Vol.6 , pp. 265-281.
- Evans, G., 1982. *The Varieties of Reference*. New York: Oxford University Press.
- Eysenck, H.J., 1995. *Genius the Natural History of Creativity*. Cambridge: Cambridge University Press.
- Feldman, D. H., Csikszentmihali, M., Gardner, H., 1994. *Changing the world: A framework for the study of creativity*. Westport, CT, Publishers/Greenwood Publishing Group.
- Feldman, R., 2000. The Ethics of Belief. *Philosophy and Phenomenological Research*, Vol. 60, pp. 671-672.
- Fischer, J. M. and Ravizza, M., 1986. *Moral Responsibility*, Ithaca: Cornell University Press.
- Fischer, J. M. and Ravizza, M., 1998. *Responsibility and Control: A theory of moral responsibility*. New York, USA: Cambridge University Press.
- Fleck, J. and Weisberg, R. W., 2004. The use of verbal protocols as data: An analysis of insight in the candle problem. *Memory & Cognition*, Vol. 32 (6), pp. 990-1006.
- Frankfurt, H. G., 1969. Alternate Possibilities and Moral Responsibility. *Journal of Philosophy*, Vol. 66 (3), pp. 829-39.
- Frankfurt, H. G., 1971. Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, Vol. 68 (Jan.), pp. 5-20.
- Galton, F., 1869. *Hereditary Genius*. New York: Macmillan.
- Gaut, B., 2010. The Philosophy of Creativity. *Philosophy Compass*, Vol. 5, Issue 12, pp. 1034–1046.
- Gaut, B., 2012. Creativity and Rationality. *Journal of Aesthetics and Art Criticism*, Vol. 70 (3), pp. 259-270.
- Gendolla, G. H. E., 2000. On the impact of mood on behavior: An integrative theory and a review. *Review of General Psychology*, Vol. 4 (4), pp. 378-408.
- Getzels, J. and Csikszentmihalyi, M., 1976. *The Creative Vision, a Longitudinal study of problem finding in art*. New York: Wiley Interscience.
- Gilhooly, K. J., Fioratou, E. and Henretty, N., 2010. Verbalization and problem solving: Insight and spatial factors. *British Journal of Psychology*, Vol. 101 (1), pp. 81-93.
- Gilhooly, K. J. and Murphy, P., 2005. Differentiating insight from non-insight problems. *Thinking & Reasoning*, Vol. 11 (3), pp. 279-302.
- Ginet, C., 1970. Can the Will be Caused? *The Philosophical Review*, 71.1 (1962): pp. 49-55.
- Ginet, C., 1990. *On Action*. Cambridge: Cambridge University Press.
- Goldman, A. I., 1971. The Individuation of Action. *Journal of Philosophy*, 68 (21), pp. 761-774.

- Götz, I. L., 1981. On Defining Creativity. *The Journal of Aesthetics and Art Criticism*, Vol. 39, pp. 297-301.
- Gruber, H. E. and Wallace, D. B., 1999. The Case Study Method and Evolving Systems Approach for Understanding Unique Creative People at Work In: R. J. Sternberg, ed. 1999. *The Handbook of Creativity*. New York: Cambridge University Press.
- Guilford, J. P., 1959. *American Psychologist*. (Aug.) Vol.14 (8), pp. 469-479.
- Helm, P., 1994. *Belief Policies*. New York: Cambridge University Press.
- Hieronimi, P., 2008. Responsibility for Believing. *Synthese*, Vol.161 (3), pp. 357-373.
- *Hieronimi, P., 2009. Two Kinds of Agency. In: L. O'Brian and M. Soteriou, eds. 2009. *Mental Actions*. New York: Oxford University Press. pp. 17-37.
- Hülshager, U. R., Anderson, N. and Salgado, J. F., 2009. Team-level predictors of innovation at work: A comprehensive meta-analysis spanning three decades of research. *Journal of Applied Psychology*, Vol. 94 (5), pp. 1128-1145.
- James, W., 1890. *The Principles of Psychology*. London: Macmillan and Co.
- James, W., 1905. The Will to Believe. In: W. James, ed. 1905. *The Will to Believe and Other Essays in Popular Philosophy*. New York, London: Longmans Green and Co. pp. 1-31.
- Jausovec, N. and Bakracevic, K., 1995. What can heart rate tell us about the creative process? *Creativity Research Journal*, Vol. 8 (1), pp. 11-24.
- Jung-Beeman, M., Bowden, E. M., 2003. Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review*, Vol. 10 (3), pp. 730-737.
- Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J. and Kounios, J., 2004. Neural activity when people solve verbal problems with insight. *PLoS Biol*, Vol. 2 (4), pp. 501-510.
- Kane, R., 1985. *Free Will and Values*. New York: State University of New York Press.
- Kane, R., 1998. *The Significance of Free Will*. New York: Oxford University Press.
- Kane, R., 2002. The contours of the contemporary free will debates. In: R. Kane, ed. 2002. *The Oxford Handbook of Free Will*, Oxford, Oxford University Press. p3.
- *Kant, I., 2010. *Critique of Judgement*. Digireads.com publishing, originally published 1790.
- Kavka, G. S., 1983. The Toxin Puzzle. *Analysis*, Vol. 43 (1), pp. 33-36.
- Kershner, J. R. and Ledger, G., 1985. Effect of sex, intelligence, and style of thinking on creativity: A comparison of gifted and average IQ children. *Journal of Personality and Social Psychology*, Vol. 48 (4), pp.1033-1040.
- Kim, K. H., 2006. Can We Trust Creativity Tests? A Review of the Torrance Tests of Creative Thinking. *Creativity Research Journal*, Vol. 18, pp. 3-14.
- Kinney, D. K., Richard, R., Lowing, P. A., Leblanc, D., Zimblist, M. E., Harlan, P., 2001. Creativity in Offspring of Schizophrenic and Control Parents: An Adoption Study. *Creativity Research Journal*, Vol. 13, pp.17-25.
- Klausen, S. H., 2010. The Notion of Creativity Revisited: A Philosophical Perspective on Creativity Research. *Creativity Research Journal*, Vol. 22, Issue 4, 2010, pp. 347-360.
- Kosgard, C. M., 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Kounios, J., Fleck, J.I, Green, D. L., Payne, L., Stevenson, L., Bowden, E. M., Jung-Beeman, M., 2008. The origins of insight in resting-state brain activity. *Neuropsychologia*, Vol. 46 (1), pp.281-291.
- Kounios, J. and Jung-Beeman, M., 2009. The Aha! Moment: The Cognitive Neuroscience of Insight. *Current Directions in Psychological Science*, Vol. 18, No. 4, pp. 210-216.
- Kozbelt, A., Beghetto, R. A and Runco, M. A., 2010. Theories of Creativity. In: J.C. Kaufman and R.J. Sternberg, eds. 2010. *The Cambridge Handbook of Creativity*. New York: Cambridge University Press. pp.20-47.
- Lavric, A., Firstmeier, S. and Rippon, G., 2000. Differences in working memory involvement in analytical and creative tasks: An ERP study. *NeuroReport: For Rapid Communication of Neuroscience Research*, Vol. 11 (8), pp. 1613-1618.
- Levinson, J. L., 2003. Elster on Artistic Creativity, In: B. Gaut and P. Livingston, eds. *The Creation of Art: New Essays in Philosophical Aesthetics*, New York: Cambridge University Press.
- Levy, K., 2009. On the Rationalist Solution to Gregory Kavka's Toxin Puzzle. *Pacific Philosophical Quarterly*, Vol. 90 (2), pp. 267-289.

- Lopez, J. C. and Lopez, D., 1985. Killer whales (*Orcinus orca*) of Patagonia, and their behavior of intentional stranding while hunting nearshore. *Journal of Mammalogy*, Vol. 66, pp. 181-183.
- Lubart, T. I., 1999. Creativity across cultures. In: R. J. Sternberg, ed. 1999. *The Handbook of Creativity*, New York: Cambridge University Press.
- Lubart, T. I., 2000-2001. Models of the Creative Process: Past, Present and Future. *Creativity Research Journal*, Vol. 13, Nos. 3&4, pp. 295-308.
- Lumsden, C. J., 1999. Evolving creative minds: Stories and mechanisms. In: R.J. Sternberg, ed. 1999. *The Handbook of Creativity*. New York: Cambridge University Press.
- Luo, J., Niki, K. and Philips, S., 2004. Neural correlates of the Aha! reaction. *NeuroReport*, Vol.15, pp. 2013-2018.
- Maier, N. R. F., 1970. *Problem Solving and Creativity; In Individuals and Groups*. California: Brooks Cole.
- Marsike, M. and Margrett, J. A., 2006. Everyday Problem Solving and Decision Making. In: J.E. Birren, K. Warner and K.S Schaie. 2006. *Handbook of the Psychology of Ageing*. 6th edition, place: Academic Press pp315-352.
- Martindale, C., 1999. 7 Biological Bases of Creativity. In: R.J. Sternberg, ed. 1999. *The Handbook of Creativity*. New York: Cambridge University Press.
- Martindale, C., Anderson, K., Moore, K. and West, A.N., 1996. Creativity, oversensitivity and rate of habituation. *Personality and Individual Differences*, Vol. 20, pp. 423-427.
- Mayer, R. E., 1999. 'Fifty Years of Creativity Research' In: R.J. Sternberg, ed. 1999. *The Handbook of Creativity*. New York: Cambridge University Press.
- McCann, H., 1974. Volition and Basic Action. *The Philosophy Review*, Vol. 83, No. 4, pp. 451-473.
- McGinn, C., 2004. *Mind Sight*. Cambridge, Massachusetts, & London, Harvard University Press.
- McHugh, C., 2011. What Do We Aim At When We Believe? *Dialectica*, Vol. 65 (3), pp. 369-392.
- Mele, A. R., 1992a. *Springs of action: Understanding Intentional Behavior*. New York: Oxford University Press.
- Mele, A. R., 1992b. Recent Work on Intentional Action. *American Philosophical Quarterly*, Vol. 29 (3), pp. 199- 217.
- Mele, A. R., 1992c. Intentions, Reasons, and Beliefs: Morals of the Toxin Puzzle. *Philosophical Studies*, Vol. 68 (2), pp. 171 - 194.
- Mele, A. R., 1997. Agency and Mental Action. *Nous*, Vol. 31, pp. 231-249.
- Mele, A. R., 2009. Mental Action: A Case Study. In: L. O'Brian and M. Soteriou, eds. 2009. *Mental Actions*. New York: Oxford University Press. pp. 17-37.
- Mele, A. R. and Moser, P. K., 1994. Intentional Action. *Nous*, Vol. 28 (1), pp. 39-68.
- Mele, A. R. and Sverdlik, S., 1996. Intention, Intentional Action, and Moral Responsibility. *Philosophical Studies*, Vol. 82 (3), pp. 265 - 287.
- Metcalf, J. and Wiebe, D., 1987. Intuition in insight and noninsight problem solving. *Memory & Cognition*, Vol. 15 (3), pp. 238-246.
- Mill, J., 1869. *Analysis of the Phenomena of the Human Mind*. (pp. 330-331). London: Longmans.
- Miller, A. J., 1992. Scientific creativity: A comparative study of Henri Poincare and Albert Einstein. *Creativity Research Journal*, Vol. 5, pp. 385-414.
- Moran, R., 2001. *Authority and Estrangement: An Essay on Self-knowledge*. Princeton, NJ: Princeton University Press.
- Mumford, M. D., Baughman, W. A., Supinski, E. P. and Maher, M. A., 1996. Process-Based Measures of Creative Problem Solving Skills: II. Information Encoding. '*Creativity Research Journal*, Vol. 9 (1), pp. 77-88.
- Mumford, M. D. and Gustafson, S. B., 1998. Creativity syndrome: Integration, application and innovation. *Psychological Bulletin*, Vol. 103 (1), Jan 1988, pp. 27-43.
- Mumford, M. D., Mobley, M. I., Uhlman, C. E., Reiter-Palmon, R., and Doares, L. M., 1991. Process Analytic Models of Creative Capacities. *Creativity Research Journal*, 4, pp. 91-122.
- Mumford, M. D., Supinski, E. P., Baughman, W. A., Costanza, D. P. and Threlfall, K. V., 1997. Process-Based Measures of Creative Problem Solving Skills v. Overall prediction. *Creativity Research Journal*, 10, pp. 73-86.
- Nickerson, R. S., 1999. Enhancing Creativity. In: R.J. Sternberg, ed. 1999. *Handbook of Creativity*. New York, NY, US: Cambridge: Cambridge University Press. pp. 392-430.
- Nietzsche, F., 1967. *The Birth of Tragedy and the Case of Wagner*, translated by Walter Kaufmann. New York: Random House.
- Nozick, B., 1973. Trying. *The Journal of Philosophy*, Vol. 70, No. 13, pp. 365-386.

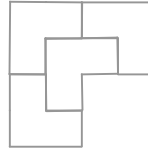
- O'Brian, L. and Soteriou, M., 2009. *Mental Actions*. New York: Oxford University Press.
- Ohlsson, S., 1984. Summary and critique of the Gestalt theory of problem solving. *Scandinavian Journal of Psychology*, Vol. 25, Issue 1, pp. 65-78.
- Okuda, S. M., Runco, M. A. and Berger, D. E., 1991. Creativity and the Finding and Solving of Real-world Problems. *Journal of Psychoeducational assessment*, Vol. 9, pp. 45-53.
- Owens, D. J., 2003. Does Belief Have an Aim? *Philosophical Studies*, Vol. 115, pp. 283-305.
- Padesky, A. and Dennis, G., 1995. *Mind Over Mood: Change How You Feel by Changing the Way You Think*. New York: Guilford Press.
- Patrick, A., 1986. The role of ability in creative "incubation". *Personality and Individual Differences* Vol. 7 (2), pp.169-174.
- Paulus, P. B., 2000. Groups, Teams and Creativity: The Creative Potential of Idea-generating Groups. *International Association for Applied Psychology*, Vol. 49, Issue 2, pp. 237-262.
- Paulus, P. B. and Nijstad, B. A., 2003. *Group Creativity: Innovation through Collaboration*. Oxford: Oxford University Press.
- Paulus, P. B. and Yang, H. C., 2000. Idea Generation in Groups: A Basis for Creativity in Organizations. *Organizational Behavior and Human Decision Processes*, Vol. 82, pp. 76-87.
- Perkins, D., 1995. Unfair review of Margaret Boden's the creative mind from the perspective of creative systems. *Artificial Intelligence*, Vol. 79, No. 1, pp. 97-109.
- Pfeiffer, R. S., 1979. The Scientific Concept of Creativity. *Educational Theory*, Vol. 29, Issue 2, pp. 129-137.
- Plato, 2004. *ION*. First published, 380bc. Translation, B. Jowett.
- Poincaré, H., 1908. In: *The Foundations of Science*. New York: Science Press of America.
- Poincaré, H., 2000. In: *Resonance*. Vol. 5, No. 2, pp. 85-94.
- Pope, R., 2005. *Creativity: Theory, History and Practice*. London: Routledge.
- Posner, M. I., 1973. *Cognition: An Introduction*. Glenview, IL: Scott Foresman.
- Pretz, J. E., Naples, A. J., Sternberg, R. J, Davidson, J. E., Recognizing, defining, and representing problems. R. J. Sternberg, ed. 2003. *The psychology of problem solving*. New York, NY, US: Cambridge University Press. pp. 3-30.
- Pritchard, D., 2007. *Epistemic Luck*, New York: Oxford University Press.
- Proust, J., 2001. A Plea for Mental Acts. *Synthese*, Vol. 129(1), pp. 105-128.
- Qiu, J., Li, H., Yang, D., Luo, Y., Li, Y., Wu, Z. and Zhang, Q., 2008. The neural basis of insight problem solving: An event-related potential study. *Brain and Cognition*, Vol. 68, pp.100-106.
- Qiu, J., Luo, Y., Wu, Z. and Zhang, Q., 2006. A further study of ERP effects of "insight" in a riddle guessing task. *Acta Psychologica Sinica*, Vol. 38, pp. 507-514.
- Ram, A., Wills, L., Domeshek, E., Neressian, N., Kolodner, J., 1995. Understanding the creative mind: a review of Margaret Boden's creative mind. *Artificial Intelligence*, Vol. 79, No. 1, pp. 111-128.
- Rawls, J., 1971. *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Reiter-Palmon, R., Mumford, M. D., O'Connor, B. J. and Runco, M. A., 1997. Problem construction and creativity: The role of ability, cue consistency and active processing. *Creativity Research Journal*, 10, pp. 9-23.
- Reiter-Palmon, R. and Robinson, E. J., 2009. Problem identification and construction: What do we know, what is the future?' *Psychology of Aesthetics, Creativity and the Arts*, Vol. 3, No 1, pp. 43-47.
- Reverberi, C., Toraldo, A., Agostini, S. and Skrap, M., 2005. Better without (lateral) frontal cortex? Insight problems solved by frontal patients. *Brain*, Vol. 128 (12), pp. 2882-2890.
- Rietzschel, E. F., Nijstad, B. A and Strobe, W., 2007. Relative accessibility of domain knowledge and creativity: The effects of knowledge activation in the quantity and originality of generated ideas. *Journal of Experimental Social Psychology*, Vol. 43, pp. 933-946.
- Ritchie, G., 2006. The transformational creativity hypothesis. *New Generation Computing*, Vol. 24, pp. 241-266.
- Ritchie, G., 2007. Some Empirical Criteria for Attributing Creativity to a Computer Program. *Minds and Machines*, Vol. 17, pp. 67-99.
- Roberts, R. C. and Wood, W. J., 2007. *Intellectual virtues: An essay in regulative epistemology*. New York: Oxford University Press.
- Rose, M., Haider, H., Weiller, C. and Büchel, C., 2002. The role of medial temporal lobe structures in implicit learning: An event-related fMRI study. *Neuron*, Vol.36, pp. 1221-1231.
- Rose, M., Haider, H., Weiller, C. and Büchel, C., 2005. Unconscious detection of implicit expectancies. *Journal of Cognitive Neuroscience*, Vol.17, pp. 918-927.

- Runco, M. A. and Sakamoto, S.O., 1996. Optimization as a guiding principle in research on creative problem solving. In: T. Helstrup, G. Kaufman and K. H. Teigen, eds. 1996. *Problem solving and cognitive processes: Essays in Honour of Kjell Raaheim*. Bergen, Norway: Fagbokforlaget Vigmostad and Bjorke. pp. 119-144.
- Sankuhler, S., Bhattacharya, J., Zak, P., 2008. Deconstructing Insight: EEG Correlates of Insightful Problem Solving. *Publication*, Vol. 3, Issue 1, p. 1459.
- Sanz, C. M., Schöning, C., Morgan, D. B., 2010. Chimpanzees prey on army ants with specialized tool set. *American Journal of Primatology*, Vol. 72, pp. 17-24.
- Sartre, J. P., 1958. *Being and Nothingness*, England: Methven and Co Ltd.
- Schank, R. C. and Foster, D. A., 1995. Engineering of creativity: a review of Boden's the Creative Mind. *Artificial Intelligence*, Vol. 79, No. 1, pp. 129-143.
- Scheerer, M., 1963. Problem-solving. *Scientific American*, Vol. 208 (4), pp.118-128.
- Schooler, J. W., Ohlsson, S. and Brooks, K., 1993. Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology*, Vol. 122 (2), pp.166-183.
- Schopenhauer, A., 1969. *The World as Will and Representation*. Vols. I and II, translated by E. F. J. Payne. New York: Dover.
- Scott, G., Leritz, L. E. and Mumford, M. D., 2004. The Effectiveness of Creativity Training: A Quantitative Review. *Creativity Research Journal*, 16, pp. 361-388.
- Searle, J. R., 1979. The Intentionality of Intention and Action. *Inquiry*, Vol. 22, pp. 253-280.
- Shah, N. and Velleman, D., 2005. Doxastic Deliberation, *The Philosophical Review*, Vol. 114, No. 4, pp. 497-534.
- Shane, S. H., Wells, R. S. and Würsig, B., 2006. Ecology, Behavior and Social Organisation of the Bottlenose Dolphin: A Review. *Marine Mammal Science*, Vol. 2, pp. 34-63.
- Shoemaker, S., 2003. Moran on Self-Knowledge. *European Journal of Philosophy*, Vol. 3 (3), pp. 391-401.
- Sidgwick, H., 1907. *The Methods of Ethics*. 7th ed. Indianapolis: Hackett. Reprint of 1907 edition.
- Siemer, M., 2009. Mood experience: 'Implications of a dispositional theory of moods. *Emotion Review*, Vol. 1(3), pp. 195-205.
- Sio, U. N. and Ormerod, T. C., 2009. Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, Vol. 135 (1), pp. 94-120.
- Slepian, M. L., Weisbuch, M., Rutchick, A. M. Newman, L. S. and Ambady, N., 2010. Shedding light on insight: Priming bright ideas. *Journal of Experimental Social Psychology*, Vol. 46 (4), pp. 696-700.
- Smith, A. M., 2005. Responsibility for Attitudes: Activity and Passivity in Mental Life, *Ethics*, Vol. 115 (2), pp. 236-271.
- Starchenko, M., Bechtereva, N., Pakhomov, S. and Mendvedev, S., 2003. Study of the brain organization of creative thinking. *Human Physiology*, Vol. 29, pp. 652-653.
- Stein, M. I., 1953. Creativity and Culture. *The Journal of Psychology: Interdisciplinary and Applied*, Vol. 36, pp. 311-322.
- Sternberg, R. J., 1999. *The Handbook of Creativity*, New York, Cambridge University Press.
- Sternberg, R. J., and Davidson, J. E., 1992. Problem solving. In: M.C. Aikin ed. *Encyclopaedia of Educational Research*, Vol. 3, pp. 1037-1045. New York: Macmillan.
- Sternberg, R. J., and Lubart, T. I., 1999. The concept of creativity: Prospects and paradigms. In: R.J. Sternberg, 1999. *The Handbook of Creativity*, New York: Cambridge University Press.
- Stokes, D., 2007. Incubated Cognition and Creativity. *Journal of Consciousness Studies*, Vol.14 (3), pp. 83-100.
- Stokes, D., 2008. A Metaphysics of Creativity. In: K. Stock & K. Thomson-Jones, eds. 2008 *New Waves in Aesthetics*:Palgrave. pp. 05-124.
- Stokes, D., 2011. Minimally Creative Thought. *Metaphilosophy*, Vol 42, pp. 548-681.
- Strawson, G., 2003. Mental Ballistics. *Proceedings of the Aristotelian Society*, Vol.103 (3), pp. 227-256.
- Strawson, P. F., 1962. Freedom and Resentment. *Proceedings of the British Academy*, Vol. 48, pp. 1-25.
- Thagard, P. and Stewart, T. C., 2011. The Aha! Experience: Creativity Through Emergent Binding in Neural Networks, *Cognitive Science*, Vol. pp. 1-33.
- Topolinski, S. and Reber, R., 2010. Gaining Insight into the "Aha" Experience. *Current Directions in Psychological Science*, Vol.19 (6), pp. 402-405.
- Toribio, J., 2011. What We Do When We Judge. *Dialectica*, Vol. 65 (3), pp. 345-367.

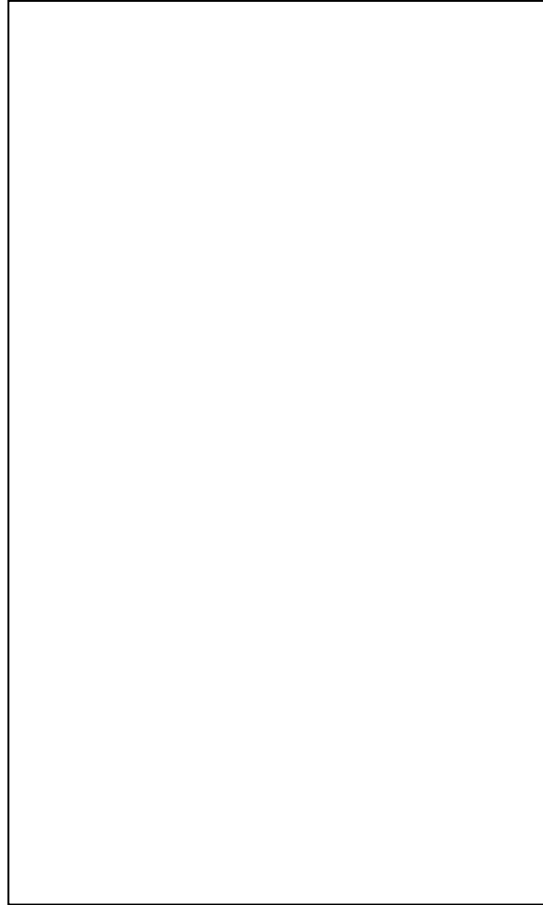
- Torrance, E. P., 1974. *Torrance Tests of Creative Thinking: Norms-technical manual*. Research Edition-Verbal Tests, Forms A and B- Figural Tests, Forms A and B. Princeton, NJ: Personnel Press.
- Torrance, E. P., 1988. The nature of creativity as manifest in its testing. In: R.J. Sternberg, ed. 1988. *The Nature of Creativity: Contemporary Psychological Perspectives*. New York: Cambridge university press. pp. 43-75.
- Turner, S. R., 1995. Margaret Boden, the creative mind. *Artificial Intelligence*, Vol. 79, No. 1, pp. 145-159.
- Tweney, R. D., 1992. Inventing the field: Michael Faraday and the creative "engineering" of electromagnetic field theory. In: R. J. Sternberg and D.N. Weber, eds. 1992. *Inventive Minds: Creativity in Technology*. Oxford . Oxford University Press. pp. 31-47.
- Unsworth, K., 2001. Unpacking Creativity. *The Academy of Management Review*, Vol. 26, pp. 289-297.
- Velleman, D. J., 1992. What Happens When Someone Acts? *Mind*, Vol. 101 (403), pp. 461-481.
- Vidal, R., 2009. Creativity and Problem Solvers. *AI and Society*, Vol. 23, No. 3, pp. 409-432.
- Vierkant, T., 2005. Owing Intentions and Moral Responsibility, *Ethical Theory and Moral Practice*, Vol. 8, No. 5, pp. 507-534.
- Wallace, D. B. and Gruber, H. E., 1989. *Creative People at Work: Twelve Cognitive Case Studies*. New York: Oxford University Press.
- Wallas, G., 1926. *The Art of Thought*. New York: Harcourt Brace.
- Watson, G., 2002. Review: Agency and Responsibility: A Common Sense Moral Psychology, *Mind*, Vol. 111, p. 444.
- Wiesberg, R. W., 1986. *Creativity, Genius and other Myths*. New York: W. H. Freeman.
- Wiesberg, R. W., 1993. *Creativity beyond the Myth of Genius*. New York: W. H. Freeman.
- Wiesberg, R. W., 2006. *Creativity: Understanding Innovation in Problem Solving, Science, Invention and the Arts*. New Jersey: John Wiley and Sons Inc.
- Wiesberg, R. W. and Alba, J. W., 1981. An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology*, Vol. 110 (2), Jun 1981, pp. 169-192.
- Wiggins, G. A., 2001. Towards a more precise characterization of creativity in AI. In C. Bento and A. Cardoso, editors, *Proceedings of the ICCBR2001 workshop on Creative Systems*.
- Williams, B., 1973. Deciding to Believe. *Problems of the Self*, Cambridge: Cambridge University Press, pp. 136-151.
- Wilson, G., 1989. *The Intentionality of Human Action: Revised and Enlarged Edition*. Stanford California: Stanford University Press.
- Wilson, G. and Shapall, S., 2012. Action. *The Stanford Encyclopedia of Philosophy, (Summer 2012 Edition)*.
- Wittgenstein, L., 1953. *Philosophical Investigations*. Translated by G.E.M. Amscombe. New York: Macmillan.
- Wittgenstein, L., 1981. *Zettel*, Basil: Blackwell.
- Wolf, S., 1993. *Freedom Within Reason*. New York: Oxford University Press.
- Yani, I. and Meyer, D. E., 1987. Activation and meta-cognition of inaccessible stored information: potential basis for incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 13 (2), pp. 187-205.
- Zagzebski, L. T., 1996. *Virtues of the Mind: An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Appendix 1

0.1 Answer to the figure problem from section 0.4



0.2 The problem from 1.5.1. How can you cut a hole in the paper that you can put your head through?

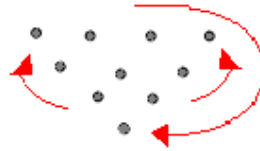


0.3 Answer to the problem from 1.5.1: Cut a spiral in the paper.

0.4 Solution to matchstick task



0.5 Solution to the triangle problem from subsection 2.3.4



0.6 The solution to the cup problem is to move glass b and pour its liquid into the glass in position E (put the empty glass back in position b).



Images in 0.1, 0.4, and 0.5 I were taken from <http://www.docstoc.com/docs/28709400/insightproblems>.