

**FACING THE MUSIC**  
**A critical enquiry into ‘the dodo-bird verdict’**  
**to develop music psychotherapy theory**

RoseAnna van Beek  
Master thesis  
Music therapy  
Faculty of Humanities  
Autumn 2015  
University of Jyväskylä

## UNIVERSITY OF JYVÄSKYLÄ

Faculty of humanities	Department of music psychology
<i>Author:</i> RoseAnna Mayflower van Beek	
<i>Title:</i> Facing the music: A critical enquiry into ‘the dodo-bird verdict’ to develop music psychotherapy theory.	
<i>Subject:</i> Music therapy	<i>Level:</i> Master thesis
December 2015	70 pages
<i>Abstract:</i> ‘The dodo-bird verdict’ represents a boggling outcome of research that shows no significant difference between the outcomes of a variety of psychotherapy treatments. Not being able to scientifically discriminate between psychotherapeutic treatments, casts doubt on their theoretical underpinnings, and even on their efficacy. This in turn presents professionals who use psychotherapeutic techniques - including music therapists - with a conundrum: What can explain this inability? And what can be done to solve the problem? Conceptual research methods were used to analyse literature in order to answer these questions. Various weaknesses were identified, both in theory construction and in research methodology within the field, which could possibly have led to the occurrence of the dodo-bird verdict. It seems that much of the conundrum could be solved by transitioning to a new theoretical framework and by incorporating new research paradigms; reimagining psychotherapy as a contextual treatment, and researching it based on systems thinking and chaos theory. Though much research remains to be done, in this new light psychotherapy - and with it music psychotherapy - stands as an effective and relevant treatment.	
<i>Keywords:</i> Music therapy, psychotherapy, contextual model, common factors, placebo effect, neurological psychotherapy, dodo-bird verdict, conceptual research.	
<i>Depository:</i> A paper copy is located at the University’s music department, a digital version in the Universities’ online database ‘JYX’ which can be found at <a href="http://jyx.jyu.fi">jyx.jyu.fi</a> .	

# CONTENTS

<b>1</b>	<b>Introducing the dodo-bird</b> .....	<b>5</b>
1.1	Music psychotherapy in the context of health and healthcare .....	5
1.2	The dilemma of choice in (music) psychotherapy treatments .....	6
1.3	Justifying trust .....	8
<b>2</b>	<b>How to tackle a dodo-bird?</b> .....	<b>9</b>
2.1	Methodology.....	9
2.2	Process description .....	12
2.2.1	Step 1: Explication and philosophical analysis .....	12
2.2.2	Step 2: Integration.....	13
2.2.3	Step 3: Reflective synthesis .....	13
2.3	Structure .....	14
<b>3</b>	<b>Origins of the dodo-bird</b> .....	<b>15</b>
3.1	Dissecting the dodo-bird verdict.....	15
3.2	Examining foundations .....	16
3.2.1	A foundation built from assumptions.....	16
3.2.2	Belief requires justification.....	17
3.2.3	Devising empirical tests .....	19
3.2.4	Applied science: an added layer of complexity .....	20
3.3	Dodo-bird verdict causes: flaws in theory construction .....	21
3.3.1	Clarification of terms .....	22
3.3.2	Foundations for pathology and treatment theories: flaws in scientific reasoning .....	22
3.3.3	Pathology and treatment theories: incorporating the humanities' perspective.....	24
3.4	Dodo-bird verdict causes: weak points in research.....	26
3.4.1	Clarification of terms .....	26
3.4.2	Problems in devising objective measurements .....	27
3.4.3	Problems in defining variables.....	31
3.4.4	Problems in controlling for confounding factors .....	33
3.4.5	Ethical limitations in designing psychotherapy research .....	35

<b>4</b>	<b>Dealing with the dodo-bird</b> .....	<b>37</b>
4.1	Addressing problems in formulating psychotherapy theories.....	37
4.1.1	A physical definition of psychological functioning.....	37
4.1.2	Qualifying health.....	40
4.1.3	Alternative ways of looking at pathology and treatment theories.....	41
4.2	Improving experimental design?.....	44
4.2.1	Exchanging linear certainty for system's based reasoning.....	45
4.2.2	A different perspective on placebos.....	46
4.2.3	The role of context in healthcare research design.....	47
<b>5</b>	<b>Beyond the dodo-bird</b> .....	<b>50</b>
5.1	Moving towards a new framework for psychotherapy.....	50
5.1.1	Contextual components in human brain development.....	51
5.1.2	The importance of human interaction.....	53
5.1.3	Psychotherapy: a valid treatment option?.....	55
5.2	Does music psychotherapy have a place within the new framework?.....	56
<b>6</b>	<b>Summary and reflection</b> .....	<b>59</b>
6.1	Research methodology.....	59
6.2	Examining foundations.....	60
6.2.1	Identifying weaknesses in theory construction.....	60
6.2.2	Identifying weaknesses in psychotherapy research.....	61
6.3	The music in music psychotherapy.....	63
6.4	Reflection.....	64
	<b>References</b> .....	<b>65</b>
	<b>List of figures and abbreviations</b> .....	<b>70</b>

# 1 INTRODUCING THE DODO-BIRD

As mentioned in the main title, the topic of this thesis concerns the so-called ‘dodo-bird verdict’, a term first introduced by Rosenzweig (1936), and its implications for music psychotherapy theory. The dodo-bird verdict (DBV) is the conclusion that can be drawn from psychotherapy research comparing different psychotherapy treatments; even though they widely vary in form and in the explanatory theories underpinning them, they all seem equally effective when tested (Rosenzweig, 1936; Luborsky, Singer & Luborski, 1975; Smith & Glass, 1977; Shapiro & Shapiro, 1982; Wampold, 2001). In essence the DBV represents a major obstacle in constructing scientifically valid mental healthcare. The necessity of scientific validity for music psychotherapy is widely recognised, often referred to as ‘evidence-based practice’ (Wigram, 2014; Otera, 2013; Vink & Bruinsma, 2003; Edwards, 2002).

The aim of this introduction is to understand what the obstacle of the DBV entails, and why overcoming it is important for progress in and credibility of music psychotherapy. There to we must first explore some of the key terms relating to the verdict; health, healthcare, psychotherapy, and music psychotherapy.

## 1.1 Music psychotherapy in the context of health and healthcare

Even though it may seem a rather straightforward concept, health is a complex thing. In this thesis a perspective from the natural sciences will be borrowed, reasons for this will be explored later on. When understanding health from this perspective, we should turn to biology and thus to the theory of evolution (Darwin, 1869). From this perspective we can understand that ‘health’ is not *one* clearly defined state that can be reached and maintained easily. Quite the opposite, it is a dynamic balancing act. It is the extent to which an organism is able to function and prosper in its locale, with the competitors and allies that surround it, and in that moment and in its particular developmental phase.

This dynamic nature of health and disease make studying them rather challenging. This is, however, the challenge that everyone who tries to promote health and development faces. In following with the definition of health proposed above, *healthcare* can be defined as the act of

facing this challenge. When difficulties in functioning negatively affect survival and prosperity of organisms we care about - most notably human beings -, finding ways to cure, prevent, or slow down degeneration of functioning becomes our goal. How should music psychotherapy be understood in the context of this definition of healthcare?

Many different definitions have been proposed, see for example Bruscia's (1998) introductory chapter in his book on music psychotherapy. Even though the definition used in this thesis does not differ much from Bruscia's example in its essence, the perspective taken is somewhat different. We shall consider music psychotherapy (MPT) to be the combination of two somewhat related fields of healthcare: Music therapy and psychotherapy.

The term music therapy (MT), so named for the treatment medium it utilises, shall in this thesis denote a healthcare treatment which uses music - perhaps among other things - to accomplish its treatment outcomes. Psychotherapy (PT), named for the type of functioning it targets, will denote any healthcare treatment targeted towards problems in psychological functioning. Quite obviously there is an overlap between the two fields, namely any type of treatment in which properties of music are used to cure, prevent or delay problems in psychological functioning; this is the definition that shall be used for MPT from here on out.

## **1.2 The dilemma of choice in (music) psychotherapy treatments**

After having defined some key terms, we can return to exploring how the DBV negatively affects progress in and credibility of MPT. Even though it may be implicit, every treatment is based on an idea of what is wrong and how to fix it - a theory. This is also the case with MPT treatments, which are based on a variety of theories; see Wigram, Pedersen & Bonde (2002) for an overview. These MPT theories often combine two components: a description of how and why music is used in treatments, and a more general theoretical basis to explain and justify the psychological treatment process. This second component is usually based on general PT theories, such as for instance psychoanalysis or behavioural therapy.

Both components of MPT theories seem vital in explaining efficacy and designing MPT treatments. However, the DBV has cropped up in PT research alone. It seems to mostly imply something about the credibility of these general PT theories, not anything specific about the

use of music in therapy. Therefore this thesis will mainly focus on this second, more general, theoretical component. Keep in mind that whenever PT is referenced to in this thesis, MPT is also implicated.

These general PT theories usually seem to contain at least two components: a pathology explanation - how the problem is thought to arise -, and a proposed treatment mechanism - a set of actions that is thought to have an effect on the problem. PT theories are represented in a multitude of competing schools of thought, which differ fundamentally in their assumptions about pathology and treatment mechanisms. Some examples of these different schools of thought are: cognitive-behaviourism, psychoanalysis and client-centered therapy. Music psychotherapists can be found representing these schools of PT, and many more.

This multitude of treatments presents a challenging puzzle to healthcare seekers and providers alike. Which PT treatment school should be chosen? And on which basis should that choice be made? There is a scientific way to answer these questions. It is first determined which treatments are effective in the first place, by measuring their effectiveness/efficacy in curing, preventing, or slowing down the degeneration of that particular functioning. Once enough different treatments for the same problem are tested in this way, it becomes possible to compare treatment effectiveness/efficacy through meta analysis. This should, at least in theory, enable a scientific answer to the challenging puzzle mentioned above.

The competing PT schools attempted to do just this. The outcome at which this type of comparison arrived, however, is puzzling. Barring the uncertainty about interpretation of the results - reasons for this will be explored later on -, comparisons have invariably found that each tested method performed equally well (Rosenzweig, 1936; Luborsky, Singer & Luborski, 1975; Smith & Glass, 1977; Shapiro & Shapiro, 1982; Wampold, 2001). This research outcome was whimsically nicknamed the 'dodo-bird verdict' (Rosenzweig, 1936), inspired by the Dodo from *Alice in Wonderland* (Carroll, 1865). The bird declared, after a very chaotic race with no clear winners, that everybody should receive a prize because they had all won. Since the different PT methods are all based on distinctly different theoretical underpinnings, this outcome seems to cast doubt either on the reliability and validity of the research, or on the accuracy of the competing theories.

### 1.3 Justifying trust

Now if you happen to be a dodo-bird organising a silly race on the beach, there is no problem with declaring everybody a winner and dolling out prizes for all. As therapists, however, people entrust their care to us. If we have the honest intention of providing them with the best possible help, resigning ourselves to this outcome seems insufficient to me. For how can one tell the difference between only winners, and only losers - between all PT methods being equally valid, and not actually knowing whether you are in fact providing adequate care?

Furthermore, apart from raising these ethical issues, the DBV situation can also negatively impact monetary compensation for PT - and rightly so I would say. Why would clients or insurance companies pay for treatments that have failed to show that they are built on solid theoretical ground? Though we may intuitively feel that our method of choice is effective, to me this feeling does not provide adequate reassurance. These tests are after all not arbitrary, as we shall see later on. They were designed to compensate for bias - our inherent human fallibility.

Learning about the DBV during my training shook me to the core, as I hope it would anyone who seriously considers the above stated facts. If we choose to acknowledge the dodo-bird verdict and what it seems to say about PT - if we face the music so to speak - can we still conscientiously offer any type of (M)PT to a client? As a beginning music psychotherapy practitioner and researcher, these questions concerned me deeply, which informed my motivation for writing a thesis on this topic. Though this core question is a hard one to answer, the next chapter delineates the research method I chose to attempt to tackle it nonetheless.



## 2 HOW TO TACKLE A DODO-BIRD?

As discussed in the introduction, the occurrence of a dodo-bird verdict when attempting to compare different types of PT treatments is reason for scepticism. The verdict casts doubt on the accuracy of the measurements used in psychotherapy research, and/or on the validity of the different theories on which the treatments are based. The issue at the core of this thesis is thus as follows: should MPT remain to be prescribed to treat problems in psychological functioning, despite the occurrence of a DBV in PT outcome research?

To work towards an answer to this core dilemma, an understanding of the possible causes of the DBV must be reached, and an overview must be made for what would need to happen to rectify the situation. To meet this aim two main research questions will be focussed on: What could have caused the DBV to occur in PT research? And what are possible ways to deal with these causes? Finding answers to these questions should allow for the formation of an opinion on whether or not MPT can remain to be conscientiously prescribed to help clients. This chapter examines ways to address the research questions, and the choice of research methodology is explained.

### 2.1 Methodology

According to Thyer (2001<sup>b</sup>) some types of research questions can be answered directly through observation or experimentation. This is called empirical research, which can use both qualitative and quantitative analysis methods. When however a large number of studies accumulate, the sheer amount of - sometimes contradictory - outcomes can obscure our understanding of a phenomenon. For this situation Thyer et al. (2001) recommend a different type of research, known as *conceptual research*.

According to Thyer et al. (2001) the goal of conceptual research is to put data into context in order to critically assess current understanding. In essence it aims to build the bigger picture. Examples of conceptual research methodologies are: theory development, historical research, literature reviews, and critical analyses. These methodologies can be used separately or in tandem with each other, and they can be combined with other (empirical) methods. Since the dodo-bird verdict is an unexpected outcome of analysing a large amount of previous studies,

conceptual research seems the type of research best suited to exploring the causes of and solutions to that situation.

When taking up research with such a wide aim some methodological challenges seem to arise. Challenges such as finding and combining the relevant sources, correctly and critically analysing them, and reformulating the findings into a useful form for the target audience. There is a snag though: there is no set methodological format for the type of conceptual research that aims to critically analyse and develop theory. How then should one go about answering these types of research questions?

Conceptual research employs secondary data. This is data collected in other - sometimes unrelated studies - as opposed to data collected specifically for the study. Because of the broad scope of the research questions finding and combining the right sources for critical analysis is perhaps the part most vulnerable to error and omissions. Even if the thesis questions are precisely stated, the causes and solutions we are attempting to find could potentially be found in any number of unlikely places in the vast amount of texts written about related subjects.

Greenhalgh and Peacock (2005) attempted to find the best solution to this challenge in a paper exploring different ways to find sources for literature reviews on complex topics. They put forward that the only way to attempt to cover such a complex topic to a satisfactory degree, is to use a search technique called *snowball sampling*. This technique combines two different ways of searching: a protocol driven search with the use of specific keywords, and a more free flowing search led by chance encounter. The sources uncovered in this first search are then used as a starting point for a more thorough search. This can for instance be done by looking at the list of sources, or by reading more work of the same authors/ within the same journals, etcetera. The authors emphasise the importance of using own knowledge and contacts in this technique as well. Some amount of uncertainty, however, still remains as to whether all relevant data was collected at the end of the study.

Since theory development and critical analysis are based mostly on logical reasoning as opposed to direct empirical testing, they are susceptible to all the normal pitfalls of human reasoning - also known as *bias*. Bias towards ideas we are already familiar with, and believe

in, or even bias against ideas which do not appeal to us for many reasons. They could include rash emotional reactions or difficulties in grasping new concepts. As studied and elegantly explained by Kahneman (2012), bias is something that pervades our thinking, it is quite difficult to avoid.

When dealing with numbers, rigorous statistics are the saving grace of scientists. When dealing with abstract concepts however, the only way through seems to be taking ones' time, reading a lot - including contradicting or otherwise unappealing source material -, and continually questioning and re-thinking what you think you know. Whether a researcher does, or does not, manage these things can seemingly only be assessed through critical reading by others.

For critically analysing texts and formulating a theoretical framework Bruscia (2012<sup>a</sup>) offers some guidelines. A theoretical text can be evaluated on the following terms: coherence, clarity, comprehensiveness, relevance and usefulness. Therefore someone attempting to critically assess a text, should keep these adjectives in mind while reflecting on what is being read. This involves making sure that the text is: internally logic - that questions, terms, and arguments are clearly described -, that the theory or text is applicable to the entire field to which it claims to apply, and that it is relevant and useful to practice. This naturally applies both to the author of a thesis, as well as to its critical readers.

Bruscia (2012<sup>b</sup>) also describes general methods that theorists can use to (re)form theoretical understanding. *Explication*: making concepts, questions, practices and terms explicit through various ways of organising and defining. *Integration*: bringing together different perspectives on the same topic. *Philosophical analysis*: exposing and evaluating underlying assumptions, and using argumentation as the primary mode of inquiry (Aigen, 2012). *Empirical analysis*: basing a theory on the analysis of empirical data. And lastly *Reflective synthesis*: the process of forming a theory through reflection on the four previous processes and on own experiences. According to Bruscia (2012<sup>b</sup>), often these methods will be combined within one theoretical text.

The thesis questions, stated in the introduction to this chapter, will be addressed using a combination of the search method put forward by Greenhalgh and Peacock (2005), and the

analysis guidelines offered by Bruscia (2012<sup>a</sup>, 2012<sup>b</sup>). For a more detailed description of what this entailed in the process of writing this thesis, see the next paragraph. As for the reporting format: according to Thyer et al. (2001) it is customary to present the result of this type of thesis in a narrative structure. Since this style seemed also to me to suit this type of thesis best, I chose to adhere to this advice.

## **2.2 Process description**

As described earlier in this chapter, the type of research in this thesis brings with it some particular reliability and validity challenges. In general I have striven to be explicit in the path that was taken to the conclusions represented. An important part of this transparency is a more in depth description of the research and writing process. The following subparagraphs paint a picture of how I went about the process of researching and writing this thesis.

### **2.2.1 Step 1: Explication and philosophical analysis**

The first step I took in this process was to take some distance from the subject. The psychotherapy theories and research which I needed to re-evaluate, had begun to feel very familiar due to my exposure to them during my education. In order to identify their underlying assumptions, I needed to be able to take a look at them with fresh eyes. I started by refreshing my understanding by reading studies and textbooks and by watching relevant educational videos (e.g. documentaries and online lectures).

During this refresher I tried to identify the underlying assumptions of the different theories and of the research itself. This led me to questions of how these assumptions could be evaluated. I found that to be able to assess the explanatory value of a theory, at least a basic understanding of general and scientific philosophy and biology - including some chemistry, neurology and evolutionary theory - are necessary. I did not have sufficient knowledge in these fields for this purpose. Therefore alongside re-examining different theories and their critiques, I attempted to gain a basic understanding in these areas. I also tried to gain understanding about what has historically stood in the way of developing psychotherapy theories to a similar degree of validity as many medical theories.

I was able to use, and had access to, both the online and physical library at the University of Jyväskylä, privately owned material, and research papers, books and lectures accessible online to the general public. English, Dutch and German language material was used.

### **2.2.2 Step 2: Integration**

After gaining better understanding of underlying assumptions on which the different psychotherapy theories were built, and what has held these theories back from being more thoroughly developed, I shifted my attention to finding research about possible solutions to the identified problems.

Making notes and sketches helped to clarify what I had learnt and it led to new questions, which sent me off into new directions of research. A great deal of the answers and new questions I found, and the ideas for new directions to search in, were the product of discussions with colleagues, lecturers, classmates, friends and family. I found my way to books and research papers by people who have been trying to solve the same puzzle - or a part of it. With the understanding I was – hopefully - gaining, I attempted to assess the usefulness of the ideas I came across to base suggestions for MPT on.

### **2.2.3 Step 3: Reflective synthesis**

Over this period, this process eventually led to a synthesis of insights on the subject, and what they could look like when applied to MPT. It was a non-linear form of research, wherein during the process the goals were not always clear, and the path taken was not always straightforward - sometimes jumping from step 1 to step 3, back to step 2, only to end up more confused.

By writing and re-writing I ordered and re-ordered my thoughts and conceptions, continually attempting to clarify points I did not understand yet. In this phase I shared my evolving ideas, sometimes only understanding a particularly part of the narrative after - successfully or not - explaining it to someone else. I have tried to make my argumentation clear and concise. To test and improve the clarity of my reasoning and argumentation, this thesis was read and critiqued by both people familiar and unfamiliar with the content matter.

To help the reader navigate and understand the text resulting from such a free flowing process, the following paragraph provides a description of the structure governing the rest of the thesis.

## **2.3 Structure**

This thesis consists of six chapters in total. The first two were an introduction to the topic and an explication of the method used to explore it. Chapters three and four focus on answering the research questions, namely attempting to answer what the likely causes for the DBV in PT research are, and how the identified problems could perhaps be remedied.

In chapter five an answer to the dilemma at the core of the thesis will be discussed - whether, when the full significance of the DBV is taken into account, there remain sufficient grounds to warrant the continued prescription of MPT to clients. A new framework for understanding PT, which implicitly emerges from the preceding chapters, will be explicated, and MPT's position within that framework will be discussed.

The final chapter is dedicated to a summary and a reflection on the process.

### 3 ORIGINS OF THE DODO-BIRD

What could have caused the dodo-bird verdict, a term first introduced by Rosenzweig (1936), to occur in PT research? To tackle this question, first the definition and nature of DBVs will be discussed, followed by an examination of the foundations on which PT theories and research are built. The last two paragraphs (3.3 and 3.4) are dedicated to discussing the possible causes of the DBV in PT research specifically.

#### 3.1 Dissecting the dodo-bird verdict

As far as I have found in the literature, the term ‘dodo-bird verdict’ has so far only been applied to the particular situation as it has occurred in PT research. However, to grasp what a DBV really *implies*, we should take a wider perspective.

In its original context the term ‘dodo-bird verdict’ specifically refers to the puzzling failure of research endeavours to distinguish between different *PT treatments*. The term, however, is a metaphor and could be understood to mean any similar type of research result. In other words, this broader perspective would define a DBV simply as a specific type of outcome: the failure of an experiment to demonstrate a significant difference between experimental conditions, even though that outcome flies in the face of how we commonly understand the world. Another way to look at a DBV is that, though it is puzzling and unsettling, it presents us with a wonderful opportunity to learn something new. Throughout the rest of this text the term DBV will refer to this new, broader definition.

A DBV conceptualised as such, is tied to a particular research method known as empirical research. Empirical research consists of formulating and performing an experiment or observation in order to find causal relationships. This causal understanding in turn is thought to provide reliable predictions for outcomes of future events similar to the experimental conditions (Thyer, 2001<sup>a</sup>). All types of research are based on some kind of epistemological philosophy - an underlying notion of what knowledge is, how it can be obtained, and what makes it reliable - or not. By examining the epistemological reasoning on which empirical research is based, different possible explanations for why an experiment can result in a DBV can be suggested.

## 3.2 Examining foundations

The concept of knowledge is less well defined than one might think. For our purpose the online Oxford dictionary's (2014) definition will be used: '[in philosophy:] *True, justified belief; certain understanding, as opposed to opinion*'. In other words, knowledge is no more than an idea that is *believed* to be true. In the case of this thesis, our beliefs about the consequences and effectiveness of healthcare procedures or tools. But what do we base our beliefs on? How should we choose which ideas to judge true or justified? Epistemology, a subfield of philosophy, is dedicated to this and other questions related to the study of knowledge.

### 3.2.1 A foundation built from assumptions

In order for any type of knowing to occur, assumptions about the nature of reality have to be made that can never be substantiated. For instance imagine the following discussion about reality. You might start with posing an observation; 'I know that I am real because when I touch my own arm, I feel resistance.' Your discussion partner may pose with; 'why would physical observations mean that you are *real*?' Coming up with an answer that in essence is better than 'well - they just do!', or 'because I said so!', is not easy. Go ahead, try. Even though it is possible to disagree on even this fundamental level, there seems no way to get out of that disagreement other than just accepting one or the other assumption to be true. (Thyer, 2001<sup>a</sup>).

As I understand it, science is an attempt to form the best possible understanding of the world around us. According to Thyer (2001<sup>a</sup>) most scientists accept a number of assumptions about reality to be self evident, because they seem to help form the most coherent picture of reality. *Realism*: that the world we observe through our senses exists independently from our mind. *Determinism*: all phenomena have physical causes which can potentially be discovered through investigation. *Positivism*: it is possible to arrive at valid knowledge about the world. *Rationalism*: reason and logic can be used to arrive at valid conclusions about observations. *Empiricism*: our senses are the only way in which we can glean original information about reality. *Parsimony*: simpler, but otherwise equally adequate, explanations take preference over



more complex ones. And *scientific scepticism*: all knowledge claims should be doubted until empirical or rational justification can be provided.

Apart from assumptions most scientists agree upon, Thyer (2001<sup>a</sup>) also lists some commonly rejected principles. *Metaphysics*: the use of non-empirical or non-rational explanations. *Nihilism*: nothing can be known or learned. *Dualism*: reality consists of two fundamentally separate parts; mind and matter. *Reification*: explaining an observation by suggesting the existence of a construct for which no valid evidence can be found. *Circular reasoning*: an explanation in which cause and effect are conflated. And *scientism*: scientific enquiry is the only valid way in which any state of 'knowing' can be reached.

### 3.2.2 Belief requires justification

Apart from agreement about which assumptions knowledge can best be built on, agreement about which type of justification is needed to call a belief *true*, is necessary as well. According to Rescher (2003) in epistemology two types of knowledge statements are generally distinguished, the distinction between the two is based on which type of justification they require to be believed: *a priori* and *a posteriori* statements.

According to Rescher (2003) *a priori* statements only require logical justification. He explains that logical justification is built on two things: an understanding of word definitions and of the rules of logic. These rules are comparable to a set of mathematical instructions that determine whether a statement is true or false. Take for instance the statement 'the person living in the house next to mine is my neighbour'. Knowing whether this statement is true requires: 1. Knowing what each word in the statement means. 2. Knowing that the logic statement 'is' means that the things before and after it are equal to each other. And 3. Understanding whether or not the words 'the person living in the house next to mine' and 'my neighbour' are by definition equivalent.

*A posteriori* statements, on the other hand, require both logical justification, as well as empirical verification, Rescher (2003) explains. The truth of this type of statement must also be checked against observations of reality. For instance knowing whether 'my neighbour's name is Stephanie' is true, requires something more. On top of the first three requirements as

described above, it requires a comparison to reality. After all, it claims something about the actual state of the world - that the person living next door to me is indeed named Stephanie.

This quirk of a posteriori statements, that their truth status can only be known *after* checking them against reality, brings us to what is known as the induction problem. Much of human prosperity - survival even - seems dependent on our ability to know whether a posteriori statements are true, already *beforehand*. For instance we rely on predictions about weather patterns for our food crops, and predictions on human behaviour for our social functioning. On which basis do we make these predictions? And how do we know whether they are worth more than lucky guesswork? Understanding the induction problem is best done through exploring an example.

Imagine living in the countryside and needing to cross a bridge every time you go into town to do your grocery shopping. You would probably like to know before crossing whether or not the bridge will carry you safely across that day. You could start by making a prediction - in research jargon this is called a hypothesis. You predict that the bridge will be safe to cross today. Next you need to check whether or not your prediction was true; you cross the bridge. If you indeed make it safely across, the confidence you have in your predictive powers will probably have increased. But next you think to yourself: 'Today was a calm, sunny day, and the bridge is new and sturdy. What if I need to cross when the wind is blowing, when it is raining, or when years have passed and the bridge's wood is starting to rot?'

You decide to make it your lifelong mission to become perfect at predicting bridge safety. You decide to test all manner of bridges, under every circumstance you can imagine. After many, many tests you feel you can safely predict whether any bridge will or will not be safe to cross. Taking past tests and applying their outcomes to new situations is called generalisation, or induction. The crux of the matter is, though, that regardless of how many tests were completed, only one counter observation is required - just one unexpected bridge collapse so to speak - to invalidate the confidence in your predictions. How then can one ever be confident about any prediction?

Popper (1972) introduced a solution to the induction problem; the epistemological system based on his ideas is called critical rationalism. The assumption underlying critical

rationalism is the acceptance that in essence we can never truly justify that an inductive statement is true or not. All theoretical knowledge that is anchored in inductive reasoning is therefore subject to its inherent flaw; any prediction (hypothesis) that follows from a theory, only requires one counter observation to prove the theory wrong. Popper's proposed solution is to form theories in a way that makes it as easy as possible to *disprove* them. He called this a falsifiable theory.

Within this epistemological system it is assumed that the only way to improve confidence in a theory's predictions, is by subjecting its hypotheses to the most rigorous empirical tests we can devise. Exactly what the bridge safety expert did in the aforementioned example. It is understood, however, that this will never lead to 'provable truth', but only to a statement that has not - yet - been falsified. The degree to which one chooses to be confident about the statement's veracity then depends on the severity of the empirical tests to which the statement has been exposed. This process should then lead to increasingly useful and robust theories that hold up under more and more rigorous experimentation. (Popper, 1972). Based on methodology textbooks on PT research - see for instance Thyer (2001<sup>b</sup>) - this stance seems to be what most empirical scientific PT research nowadays is based on.

### **3.2.3 Devising empirical tests**

So in summary, from the perspective of most scientists, a theory derives its credibility from adhering to a certain set of basic assumptions, from being falsifiably formulated and by withstanding the most rigorous empirical tests we can devise. In this last condition lies a final challenge to a theory's credibility; devising a proper empirical test can be quite complicated. Empirical research is an attempt to formally address this challenge.

According to Thyer (2001<sup>b</sup>) there are two qualifiers that can indicate how well the outcome of empirical research can be relied upon: *validity* – how well it measured what it meant to measure –, and *reliability* – the accuracy of the measurements. Many things can stand in the way of achieving validity and reliability within research, and we may not always be aware of them or be able to remedy the situation.

The observations, or measurements, that make up our tests, rely on human sensory capabilities. According to Chabris and Simons (2009) it is not very difficult to find examples of how our own senses deceive us. Furthermore, the way in which tests are devised and how results are interpreted relies on our cognitive capabilities. As Kahneman (2012) showed, human beings also do not find it hard to make errors in applying logic and interpreting information.

In an attempt to best address these challenges, empirical research follows formalised steps to mitigate the chance of observation or cognitive errors. The precise steps are too numerous and intricate to discuss here, but in general they encompass the following actions. First an explanatory theory is formulated in a falsifiable way, in order to make testing it, both logically and empirically, possible. This is done by defining all terms in a precise and measurable way. Then a prediction is formulated that describes what would be observed if the theory were false - this is called a null-hypothesis. The empirical testing is done in the form of an experiment that attempts to measure, as validly and reliably as can be achieved, whether the theory holds up or not. Subsequently the null-hypothesis is either rejected or confirmed, which should either result in revision of the theory, revision of research methods, or in added confidence in the theory's truthfulness. (Thyer, 2001<sup>b</sup>.)

### **3.2.4 Applied science: an added layer of complexity**

This paragraph has so far discussed scientific research in general. The goal of science in its most basic form is to discover knowledge purely for the sake of understanding. Healthcare research, and in particular PT research, diverges from this type of science in that it seems to have a different goal; to *apply* this knowledge to meet human needs. I would argue that this makes PT an *applied* science - just like medicine, architecture, engineering etcetera. This adds a layer of complexity to the search for DBV causes, and should be considered as well.

Being an applied science places PT research in the middle of a complicated intersection between the sciences and the humanities, because this type of science has two different kinds of puzzles to solve. The first is *how* human needs can best be met, which can be considered an empirical question. The sciences seem best equipped to answer this. However, an applied science also requires a definition of *what* these 'human needs' constitute. Since the answer to

this kind of question is dependent on the perspective one takes - long term vs. short term, collective vs. individual interests, cultural values etc. -, it requires a subjective answer. The humanities seem much better equipped to provide such an answer.

This interplay between science and the humanities can be recognised in how research for an applied science such as PT is done. It differs from the basic sciences in that it does not just aim to see how well the formulated theories agree with observations of reality. Instead theories are formulated and tested to see how well they succeed in *explaining* and *solving* the subjectively defined human problem. In the applied health sciences these explanations and solutions are referred to respectively as pathology explanations and treatments. These are tested by field specific variations on empirical research - called RCTs - meant to either isolate causes of pathology, and/or to test treatment efficacy/effectiveness.

### **3.3 Dodo-bird verdict causes: flaws in theory construction**

As was introduced at the beginning of the chapter, we are attempting to figure out what could have led to the DBV in PT research. We took a broader perspective, and noted that a DBV could be seen as the puzzling outcome of any type of empirical research: The counterintuitive outcome of an experiment in which two or more experimental conditions were compared and found equal. This is exactly what occurred while researching PT treatments: comparisons between PT treatments based on different theories have found that each tested method performs equally well (Rosenzweig, 1936; Luborsky, Singer & Luborski, 1975; Smith & Glass, 1977; Shapiro & Shapiro, 1982; Wampold, 2001).

What could have happened to produce this puzzling outcome? When following the logic set out in the preceding paragraph (3.2), two distinct possibilities come to mind. The result could be correct, which would imply that - at least some parts of - PT theories are flawed. Ambiguities in definitions, vagueness in theory formulation, as well as non-adherence to the basic epistemological principles could all have led to theory failure. However, it could also be a false negative. The implication of this would be that the theories are - at least partway - correct, but that mistakes were made in testing them empirically. Measurements could have been performed incorrectly, or observations could have been misinterpreted.

An in depth examination of the specific situation in PT theory and research is needed to discover whether one, the other, or both scenarios (in part) could have caused the DBV to occur. This paragraph (3.3) is dedicated to examining the building blocks from which PT theories are constructed. It will consider the extent to which the theory is falsifiable and whether grounding epistemological principles are being adhered to. The next paragraph (3.4) will consider the state of PT research more thoroughly.

### **3.3.1 Clarification of terms**

Let us take as a starting point the concept of health - psychological health to be exact -, and PT's aim in improving it. Health will hereby be defined as 'optimal functioning' (see also paragraph 1.1), and PT is here understood to be an applied science that aims to improve it. Based on the reasoning put forward in the paragraph on the foundations of science and applied science (3.2), this implies something about the building blocks that PT theories should be constructed from.

Applied science theories that attempt to explain health problems (pathology) and suggest solutions (treatments) need to be based on a marriage of two considerations. The first is a falsifiable - physically measurable - construct, based on scientific reasoning. The second is a definition of what 'optimal functioning' denotes, based on ideas from the humanities.

Our quest is to uncover possible flaws in PT theory that could have caused the DBV. During the research phase in which I attempted to gain a grasp on these possible causes, I uncovered some weak point in the theory that originated in both the above mentioned domains - within the parts of the theories built on scientific reasoning, and in the parts grounded within the humanities. In the following subparagraphs (3.3.2 and 3.3.3) weak point in PT theories based on both domains will be discussed. First the current and historical pathology explanations in PT research as based on scientific reasoning will be discussed.

### **3.3.2 Foundations for pathology and treatment theories: flaws in scientific reasoning**

According to Millon and Simonsen (2010), human kind's first attempts at pinpointing an origin for psychological problems - and other types of problems as well - would have been metaphysical (e.g. the supernatural/gods), or dualistic (concerning the spirit or soul as non-

embodied entity) in nature. According to Cozolino (2010) during the 19<sup>th</sup> century a shift occurred in this way of thinking.

Darwin (1869) in his treatise on evolution already spoke of the dream that one day psychology would be based on a biological understanding of the human being. New technological possibilities for exploring the microscopic make-up of the brain at the end of the 19<sup>th</sup> century, and observations that certain brain injuries could lead to very specific psychological symptoms, led to some scientists considering whether impediments in brain functioning could be the origin of psychological symptoms (Cozolino, 2010). However contrary to hopeful wishes and intentions to base psychopathology notions on the brain, the still limited technical capabilities for exploring brain functioning further, made it impossible to suggest falsifiable explanations for psychological symptoms at that time (Freud, 1968).

This in essence started two related but - until recently - impossible to merge fields of healthcare (Cozolino, 2010); neurological medicine which focuses mostly on studying the brain physically to find pathology explanations and treatment options based on its physical properties, and PT which focuses more on reported experiences and observed behaviours to understand and treat psychological functioning. The neurological side of the divide could use falsifiable and measureable constructs to base their research on, they were however incapable of suggesting brain based explanations for the more subtle problems in experiences, behaviour and pathology. This meant that they could not yet suggest treatment options based on physical interventions in the brain for all problems in psychological functioning.

Even though it was not possible yet to formulate pathology and treatment notions for every psychological problem on falsifiable definitions, the need for treatment was present. This is seemingly the void that PT theoreticians attempted to fill. This void necessitated the suggestion of a construct from which psychological symptoms were thought to originate and could be treated; a *psyche*. However, due to the technical limitations discussed earlier no objective empirical proof could - yet - be given for the existence or make-up of such a construct. This is a perfect illustration of reification (see paragraph 3.2.1).

The proposed original construct of the psyche from the early days of psychoanalysis is a good examples of this; that the mind contains three parts, the Id, Ego, and Superego (Freud, 1923) -

constructs built by using seemingly fitting metaphors, not built from a set of physically measurable characteristics. These early psyche and psychopathology constructs seem to have sprouted into many more new constructs - one for each new school of PT to be exact. This is, along with the occurrence of a DBV, a logical outcome of the situation. For when disagreement about a theory cannot be settled with empirical measurements in research - because there is nothing to measure when reified constructs are used - no competing theories can be discarded.

During the past century our knowledge about brain functioning has increased dramatically, but not until recently has it become possible to start suggesting and testing psychopathology and treatment theories that are truly based on physical constructs - on brain functioning (Cozolino, 2010). The good news is that many constructs and concepts first suggested by different PT theories do seem to have some form of physical substrate in the brain (see for instance: Berlin, 2011).

Now that merging knowledge from neurological medicine and psychotherapy is becoming possible, the question arises whether one does not make the other obsolete. In other words, now that it is becoming possible to situate explanations of psychopathology and treatment mechanisms within the physical characteristics of the brain, it perhaps places the treatment of psychopathology in the domain of neurological medicine. Does PT still have added value to offer brain treatment? We will come back to this question in chapters four and five.

### **3.3.3 Pathology and treatment theories: incorporating the humanities' perspective**

PT theories are in way nothing more than a posteriori statements (see section 3.2.2) about how psychological functioning can be improved. As with all a posteriori statements, testing their veracity involves - among other steps - comprehension of concepts the theory is comprised of, and the performance of an empirical test. Similar to how the use of reified concepts has thwarted attempts to empirically test PT theories (see 3.3.2), so do ambiguities in concept definitions defeat clear comprehension of the theory components. These ambiguities can contribute to the occurrence of a DBV by enabling errors in the process of testing a theory.



The way health is often defined within PT research manages to bring about ambiguity within PT theories. This is the case because defining pathology requires a - perhaps implicit - understanding of what the absence of pathology would look like; what the desired state of functioning is. Since only then would you know if someone is, or is not, sick. As the term suggests, what a 'desired state' is, is dependent on the perspective one takes, on subjective preferences. Can any state be deemed desirable unequivocally? To me the answer seems 'no', so let us first discuss why an objective answer to this question seems impossible.

The first obstacle to a univocal answer is that it seems unclear who should decide. Should it be a majority vote, a completely individual decision, a decision made by one wise and/or powerful individual, or by a learned panel? As can be seen in politics, none of these systems give the assurance that a satisfactory answer can be reached for all. Fundamentally this question boils down to the question that moral philosophers have puzzled over - and not found definitive answers for - for millennia (Beauchamp, 2001); is there anything that can be called intrinsically 'good'?

Furthermore, on what basis should the decision be made? Should the individual or the collective, the long or the short term interest be given precedence? Put in vernacular; 'every pro has a con'. Examples of which are the vulnerable balance between individual rights and the fight against criminality, or the tension between the need for economic growth now versus a sustainable ecology later.

Lastly, cultural values change across time and place. What is considered acceptable in one time or place, may not translate well to the next. Think for example about the difference between how homosexual behaviour was regarded in many western states fifty years ago, and how the tides are changing. Or how in psychology for a long time the absence of disturbing symptoms was considered enough to constitute a healthy state, whereas for instance the field of positive psychology now deems the experience of happiness a necessary condition too.

PT research often employs classification systems - such as the DSM (American Psychiatric Association, 2015) and the ICD/ICF (World Health Organisation, 2015) - to define pathology. Within these classification systems the subjective perspectives on which their categories are built, are not explicitly stated. It seems to me, however, that they must implicitly still be based

on decisions about these - and perhaps other - dilemmas. It is therefore no surprise that they receive much criticism, not only for their methodological shortcomings, but also because it is impossible to please everyone. Kutchins and Kirk (2003) for instance describe the history of the DSM and how according to them it is heavily influenced by government and corporate interests, and geared towards short-term capitalistic goals instead of long-term goals supporting individual health, which they deem more important.

Classification systems are necessarily based on subjective choices about health definitions. Keeping these choices implicit makes it difficult for users of these systems to discuss their merits - and also perhaps difficult to change them when necessary. More relevant to this thesis, however, is that it also allows for ambiguities to creep into theory formation and research interpretation; creating good breeding grounds for dodo-birds.

### **3.4 Dodo-bird verdict causes: weak points in research**

In the preceding paragraph it was discussed what role flaws in the theoretical foundations of PT theory could have had on reaching a DBV. In this paragraph we shall examine how PT research methods may have contributed to this as well.

#### **3.4.1 Clarification of terms**

A pathology theory is in essence an explanation of what *causes* certain problems in human functioning. Similarly a treatment theory is an explanation of how positive changes to that situation can be *caused*. Understanding causality is in essence the same as the ability to make accurate predictions; *if A, then B*. As was discussed in the second paragraph (3.2) of this chapter, to make reliable assertions about causality, the careful design and execution of an experiment is needed. The how and why of experimental design in healthcare research, has been covered thoroughly in other places, see for instance Thyer (2001<sup>b</sup>), it will therefore not be delved into deeply here. For this paragraph, however, a short description of the reasoning behind the research and the terms I will be using seems in order.

Much healthcare research takes the form of ‘randomised controlled trials’, from this point on referred to as RCTs. The design of the experiment looks somewhat like the representation of

the pair of scales in figure 1. The two sides of the scales represent two groups of test subjects that are made as equal as possible through randomly assigning people to either of the groups. They are often tested on a specific score before the trial, referred to as the dependent variable. Then both groups undergo a treatment situation where only one thing is different between the two groups, known as the independent variable. (Thyer, 2001<sup>b</sup>).

The same test is applied after the trial. If the only difference between the groups was the independent variable, then a causal link can be shown between the administration of the independent variable and the outcome. To be sure that the independent variable was truly the only difference between the before and after measurements, anything that can affect the measurement - called potential confounding factors - is made equal between the groups as well, this is called controlling. (Thyer, 2001<sup>b</sup>).

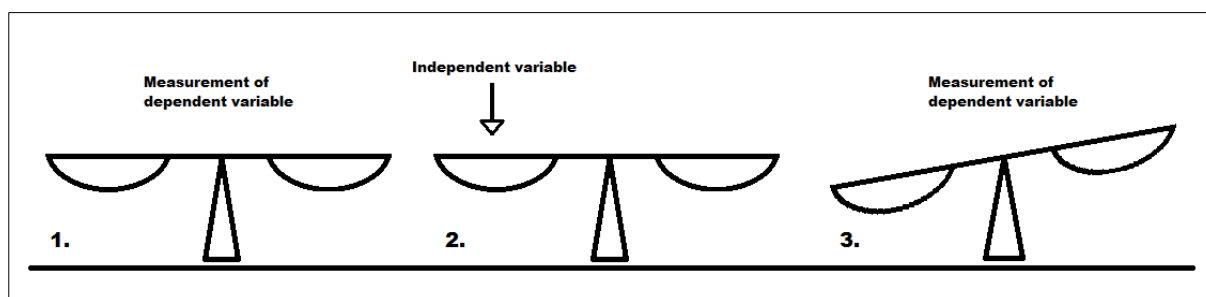


Figure 1: An illustration of RCT design.

This setup is exactly what has been attempted over the last decades with the introduction of RCTs to the PT research field (Budd & Hughes, 2009). But since this has concluded in an unexpected DBV when comparing the different PT treatments, there might be something amiss with the experimental setup. Weak points in the theory's formulation were already discussed in the previous paragraph, in this section weak points in the research design will be discussed.

### 3.4.2 Problems in devising objective measurements

To make assertions about the outcome of an experiment, observations must be made that either lead to confirmation or rejection of the null-hypothesis. Making observations however already carries its own inherent difficulty. Our sensory perceptions can for example lack precision and can be biased or flawed (Kahneman, 2012; Chabris & Simons, 2009). An

example of one such sensory limitation is that, of all the existing electromagnetic wavelengths, our human eyes can only detect a small portion - called the visible spectrum (Waldman, 2002). How then can we know to what extent any measurement, for instance PT treatment efficacy measurements, are both reliable - are accurate - and valid - measure what we mean to measure?

Firstly let us take a look at how we try to make our measurements reliable. We can attempt to detect the inherent flaws in our perceptions and devise ways to counteract them. We can for instance make the observation *indirectly*. We can invent measurement tools that can receive more information, and to a higher degree of detail and precision, than our own senses - like a microscope or a video camera. In essence use technology to translate between things we cannot sense directly, to something we can - for example false colouring or enlargement of pictures, to show what something would look like if we were able to see other wavelengths of light or discern smaller items. Simply put, according to Trochim (2006<sup>a</sup>) one can tell if a tool is reliable by persistently measuring something and getting the same result over and over again.

And what about validity? How can one tell whether the thing you are attempting to measure is actually measured by your observations or measurement tools? According to Trochim (2006<sup>b</sup>) the question in essence boils down to two types of validity; translation validity and criterion-related validity. Translation validity is concerned with to which degree a measurement tool equates to the definition of a construct. Criterion-related validity on the other hand is concerned with to what extent the measurements taken by the measurement tool correlates to the measurements predicted by the construct related theory.

If a mountain is defined as a landmass higher than a certain number of meters above sea level, a measurement tool for mountains would have translation validity if it indeed manages to distinguish between an object that does, and one that does not meet these requirements. If you have a theory concerning mountain formation, and your measurement tool manages to detect mountains in places where your theory predicts they should be, and no mountains where they shouldn't be, then your tool has criterion-related validity.

With constructs that have clear physical definitions, and when our technical capabilities are such that we can indeed measure their physical parameters - as is the case with the mountain example - both validity and reliability can be easy enough to ascertain. If however there is no clear physical definition and/or our technological capabilities are not advanced enough to measure its parameters, this can be difficult.

You may imagine that you can get around these obstacles by using indirect measurements. If you for instance want to know the average length of a male human hand, but you cannot measure hands directly for whatever reason, perhaps a measurement of a foot can be taken instead. The *assumption* is that larger individuals may have both large hands as well as large feet. To know whether such an indirect measurement can indeed stand in for a direct measurement, however, the degree of correlation must somehow be verified with direct observations. You can only be sure whether big feet do indeed equal big hands, if you measured and calculated the extent of correlation between the two in a representative sample of a large enough group. (Thyer, 2001<sup>b</sup>).

In PT research measurement tools are also used to attempt improving the reliability and validity of measurements. There are, however, some problems with these tools that make the degree to which they are valid and reliable difficult to ascertain. Even though volumes have been filled about improving the reliability and validity of psychometrics - as these measurements are called (see for instance Furr & Bacharach, 2013) -, no statistical cleverness can truly make up for some of the obstacles in our current modes of measurement. To see why we shall take a look at two main forms of measurements taken in PT outcome research: observations done by others, and self reports done by research subjects (Kaplan & Saccuzzo, 2012). In what way can these tools designed to satisfy reliability and validity criteria, and where do they fail?

Let us first discuss observations done by others. Within PT research this usually refers to observations about physical movement and verbal expressions, but in MPT this can also refer to observations about a client's musical expression. PT researchers often try to make these observations more reliable by implementing ways in which the observations become more precise and repeatable: by careful operationalisation of the observation task, and by use of technology that can record the outwardly observable variable (sound/video for instance).

There still are, however, both validity and reliability problems with this type of measurement/observation. (Kaplan & Saccuzzo, 2012).

If the psyche is assumed to be a part of, or an emergent property of, the brain (see paragraph 3.3.2), making observations of outwardly visible/audible behaviour is inherently an indirect measurement. As was discussed earlier in the example of indirect measurement of hand sizes, there is only one way to know whether an indirect measurement is a valid stand-in for a direct measurement of two variables. Only with a direct measurements of both variables (the feet and the hands in the example) can it be checked how well the direct and indirect measurement values correlate in reality, and thus whether the indirect measurement is a valid stand-in. The problem with indirect measurements of the psyche could in theory be remedied by correlating the behavioural observations to measurements of brain functioning itself. Precisely herein lies the problem; we currently do not have the capability - yet - to accurately couple behavioural observations to brain functioning; this has two reasons.

The first is that according to Fachner and Stegemann (2013) the technology that could be used to measure things about the brain in action, is not yet able to take measurements with enough accuracy within naturalistic settings. They are not yet portable enough - since they can often only be used after extensive preparation and with the subject sitting absolutely still. They are also not yet accurate enough in measuring both temporal and spatial features of activity simultaneously. The other reason why it remains too difficult to couple these observations to brain functioning was discussed already earlier already (in paragraph 3.3). As long as theories about the psyche - about the brain's social and emotional functioning - are not based on explicitly (and physically) defined concepts, there *is* nothing to couple our observations *to*.

With self-reported observations by subjects the problem is just the other way around. Since the subject is by definition the only one who has direct access to their own experiences, in a way their internal observations must by definition be valid. Even here, however, the validity and reliability of the *report* can still be questioned. When someone reports anything about internal perceptions, there is *by definition* no way to improve the reliability of these subjective observations by making them repeatable, thus the subject may be mistaken or not telling the truth and a researcher could never know. Furthermore, as was the case with behavioural observation, it is not yet possible to make sure that the subject's report is indeed coupled to

the exact phenomenon the researcher intends to study. In a way it is once again an indirect measurement, with the same pitfalls as discussed above.

### **3.4.3 Problems in defining variables**

When carrying out an RCT it is necessary to define the setting of the experiment and what is meant to be gained by it precisely. A part of this means defining what the dependent (outcome) and independent (difference between experimental and control group) variables are. Budd and Hughes (2009) identify problems in using the constructs currently used within PT research, namely DSM diagnoses and PT treatments, as dependent or independent variables. According to the authors these constructs are not precise enough to infer causal links from the studies, and they imply that this is one of the reasons for the occurrence of a DBV.

#### **DSM diagnosis as variable**

Let us begin by taking a look at DSM diagnoses as variables in an RCT. A diagnosis, as meant in this context, is the descriptor given to a specific problem in functioning. For instance the diagnosis 'flu' can be given when someone is infected with a strain of the influenza virus. However not all diagnoses are on equal footing. Some diagnoses, like the above mentioned example, are based on relatively well understood pathology mechanisms (etiology), like a viral infection. Others lack a physical explanation of what causes the symptoms, and are simply a name given to a grouping of symptoms that are thought to be linked in some way. This type of diagnosis is called a 'syndrome'. (Schneider & Lilienfeld, 2015).

In PT research a diagnosis can either be an independent variable when trying to research disease etiology, or a dependent variable when performing outcome research. So for our purposes we need to understand why the diagnoses that are currently being used in PT outcome research (the type of research that led to the DBV) are as of yet unfit to be used as dependent variables.

In PT research the diagnoses often used are derived from the DSM, a manual listing different psychological syndromes and how to recognise them (American Psychiatric Association, 2014). According to many criticisms of the DSM (for instance: Kutchins & Kirk, 2003;

Smolik, 1999) the classification system used in the DSM leads to syndromal groupings that overlap (high co-morbidity) and could easily have many different causes.

The DSM classification of diseases is not based on a system that differentiates them based on a physical nosology, but instead is based on symptom groupings that are decided on by a panel selected by the American Psychiatric Association (Kutchins & Kirk, 2003). Although there seems - as of yet - to be no better option, because our understanding of brain functioning is still too limited, it seems understandable that the arbitrary nature of this type of classification can help lead to confusing situations; as for instance a DBV.

Put in contrast to our understanding of how influenza infections occur; we simply do not know yet what exactly causes the different ‘diseases’ that are differentiated in the DSM. Some of them may in fact share causes and simply manifest differently, some may outwardly look similar but in fact not share causes at all. This means that whether changes in DSM diagnoses are measured or not, this cannot be ascribed to the treatment being tested, precluding any conclusions about causality.

### **Psychotherapy treatment as variable**

Now let us consider why using PT treatments as variables in an RCT can be problematic as well. When trying to differentiate causation from mere correlation by RCT, a researcher usually attempts to measure the effect of *one* independent variable on *one* dependent variable, for instance by measuring the effect on the functioning of a biological system after the introduction of a specific chemical compound. This is ideally done by comparing two otherwise identical circumstances; one in which that compound is introduced, and one in which it is not. Any difference in outcome between the two situations should be attributable to the independent variable (see for clarification paragraph 3.4.1).

But what happens if more than just one thing is different between the two circumstances? Making claims about causality becomes unreliable. According to Budd and Hughes (2009) a PT treatment usually takes multiple sessions, and is often performed by different therapists within one research project. They assert that, even if a treatment manual is used and the



adherence is checked by other observers, too many things can deviate between sessions and therapists to make accurate claims about causality at the end of the trial.

Another problem is that there appears to be a significant overlap between different types of PT treatment in terms of their structural features. Research that examined PT efficacy trials showed that differences between outcome measurements largely disappeared when the control groups were made structurally similar (Wampold, 2001; Baskin et al., 2003). This can be interpreted to mean that even though the different PT treatments are often thought of as separate independent variables, the structural overlap between them may in fact point to the different treatments being more similar than the different PT schools may wish to admit.

This idea that therapy efficacy may in fact arise from components of a therapy that are not acknowledged to be 'PT school specific' is often referred to as the 'common factors theory' (see for instance: Rosenzweig, 1936; Frank, 1973). If the efficacy of these different treatments lies in the treatment components they have in common as opposed to within those that differ, this would also naturally lead to the verification of the null-hypothesis (DBV) when comparing these different - but after all not so different - treatments to each other.

#### **3.4.4 Problems in controlling for confounding factors**

As described earlier (3.4.1), the way in which RCTs manage to make claims about causality, is by comparing situations in which everything except one specific variable differs between the situations. This comparison can only happen if, apart from the independent variable, everything else is the same across the two situations. This can be referred to as controlling for confounding factors. Anything that can possibly have an effect on the outcome of an experiment should be considered a potential confounding factor.

In PT research two things can be identified that have perhaps confounded research outcomes. The first, according to Budd and Hughes (2009), is the difficulty in forming homogenous test groups for PT trials, which can lead to unreliable research conclusions. They suggest that DSM diagnoses are not only used as outcome measures in PT research, but are often also used as selection criteria in trying to form homogenous testing groups. The problem with this is comparable to what was discussed in the previous subparagraph (3.4.3). If these diagnoses are

not based on valid etiological constructs, the homogeneity of the group may only be imaginary. The different groups may still be made equal through good use of randomisation, but making a reliable statement about the causation of efficacy outcomes for particular groups is in this case not possible.

The second confounding factor is somewhat harder to characterise. Budd and Hughes (2009) describe it as inaccurately using ‘the drug metaphor’, Wampold (2001) labels it the misguided attempt at ascribing the source of treatment efficacy to specific factors as opposed to contextual or common ones, and if we follow Kradin’s (2008) reasoning it has something to do with misunderstandings about the so-called placebo effect. To find clarity let us take a step back from the situation. How would one go about identifying possible confounding factors within a research design?

It seems to me that one first needs to grasp which components within the experimental setup have to potential to affect the outcome. Say for instance that you want to measure the effect of a certain sound frequency on a piece of glass. Other sounds in the same environment, or filth on the glass could perhaps distort the measured effect. They should therefore either be eliminated or controlled for in the comparison situation.

What Budd and Hughes (2009), Wampold (2001) and Kradin (2008) seem to imply, is that in PT research there is a major factor in the experimental setup that can exert influence over the outcome measure, but one that is often overlooked or misunderstood. It therefore is likely not controlled for properly. If we indeed assume that the psyche is part of, or a function of, the brain, then the question becomes as follows: which factors in PT research setups have the potential to affect brain functioning? And were they all sufficiently taken into account in PT research?

According to de Craen, Kaptchuk, Tijssen and Kleijnen (1999) observations from medical practice that something other than direct physical interventions - such as surgery or medicinal substances - seemed to be capable of affecting patient’s health outcomes, eventually led to the suggested existence of ‘placebo effects’. Budd and Hughes (2009) chronicle that when RCTs became standard procedure for testing a treatment’s efficacy, the way to control for these

effects became to add a ‘placebo control’ to the study design. But what exactly is being controlled for by adding a placebo?

Kradin (2008), Frenkel (2008), Freeman (2004), Benedetti et al (2003) and Cozolino (2010) all suggest mechanisms - largely similar ones - by which the body’s functioning can directly be influenced by perceptions from the environment, consciously or not. In essence the meaning ascribed to perceived information from the person’s context by the brain, is capable of affecting many processes within the body, including but not limited to psychological functioning. Therefore all differences in the setting of an experiment that can alter the perceptions of a test subject, have the potential to affect the outcome measures. A placebo control is therefore a way to control for ‘contextual factors’ as Wampold (2001) would call them, the ‘structural components of a treatment’ according to Baskin et al. (2003), or for ‘meaning responses’ as coined by Moerman and Jonas (2002).

The quality of placebo controls for these structural components used in PT research however varies much between different trials. As already mentioned in the previous subparagraph, Baskin et al. (2003) performed a meta-analysis of the effect of structural differences between treatment and placebo arms of RCTs. They concluded that the measured size of treatment effects was significantly correlated with the amount of structural similarity between the two arms. The smaller the difference was, the smaller the treatment effect that was measured. This again points to the necessity to conclude that the components that make PT treatments efficacious should be perhaps be sought in the components they have in common, instead of those that differ. This misconception about what should be controlled for in a PT RCT, if not addressed, would inevitably lead to a DBV.

### **3.4.5 Ethical limitations in designing psychotherapy research**

Even if perfect measurement tools existed, a flawless definition of the applicable variables were given, and if it were possible to control for every possible confounding variable, a final hurdle would still exist; ethical restrictions in testing on human (or animal) subjects. To put it differently; even if we knew how to unequivocally determine which treatment is most effective for certain psychological health problems, we may still decide that the road to this knowledge is an unacceptable one to walk.

Nowadays when testing any treatment or healthcare theory on humans (and also on animals in many cases) a board of ethics needs to be asked for permission first. Under these circumstances it can be especially difficult to adequately control for confounding factors. Getting permission when for instance subjects are not informed completely about the procedure they will be undergoing, when withholding a treatment already known to be effective, or when not enough is known yet about the risks of a certain procedure, the chances of getting permission are slim.

For safety and fairness this is undoubtedly a good thing. For our surety about the effectiveness of healthcare procedures however, it is not. Apart from all of the before discussed obstacles, it may still be impossible to test PT treatments rigorously enough to truly rule out a DBV because of the complicated ethics concerned.

## **4 DEALING WITH THE DODO-BIRD**

The previous chapter was dedicated to answering the first of two research questions. It was discussed how some of the requirements necessary to reach dependable conclusions with scientific research were - out of necessity it seems - not followed within PT research. This seems to adequately explain how the DBV could have occurred there. It feels appropriate, since it involves applied science research (see 3.3.3), that this thesis also aims to transform this new found understanding into practical enhancement of therapy and research practices. This chapter will therefore focus on answering the second research question; it will offer some solutions or encouraging avenues which could be used to deal with the DBV.

### **4.1 Addressing problems in formulating psychotherapy theories**

As discussed earlier (paragraph 3.2 and 3.4), the basis of a scientific experiment is the formulation of an idea about how the world works, and a formulation of testable predictions based on that idea; a theory with falsifiable hypotheses. On top of that, for an applied science such as PT the goal of the discipline, i.e. the human requirement, needs to be qualified so that the degree of achievement (or failure) of meeting the goal can be tested, and the goal's merit can be debated.

If the terms/constructs that make up the theory are not both falsifiably defined and explicitly characterised, it cannot become clear whether different researchers are measuring the same thing. This is where many problems in PT theory seem to stem from; difficulties with unambiguously or transparently defining the concepts on which the explanatory theories are based (see also paragraph 3.3). In this paragraph some suggestions will be discussed for improving on some basic theoretical components: a falsifiable definition of psychological functioning, a transparent qualification of psychological health, and new way to view psychopathology and treatment mechanisms.

#### **4.1.1 A physical definition of psychological functioning**

To make discussing this subject somewhat easier, let us begin with a picture to illustrate the dilemma we face when attempting to physically define psychological functioning. Figure 2

shows a simplified illustration of the puzzle; a schematic representation of psychological functioning. The input referred to in the illustration is comprised of all chemical and electrical sensory input that reaches the brain from both inside (e.g. pain, emotions) and outside the body (e.g. visual or auditory stimuli). The output represents actions actuated by the brain, implemented by the body outside or within itself in the form of behaviours (motor functioning) or physical changes (activation or inhibition of other organ functions).

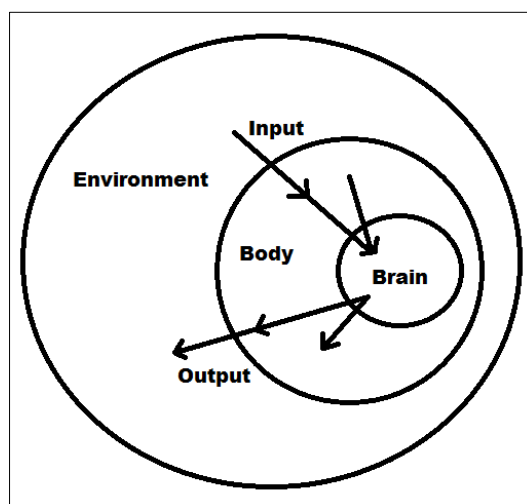


Figure 2: A schematic representation of psychological functioning.

As discussed earlier (3.3.2), the idea that the brain is responsible for the translation between input and output, as represented in figure 2, has been around since as early as end 19<sup>th</sup> century. Our physical knowledge about the brain, however, has lagged behind. This has precluded suggesting any falsifiable construct to explain how the brain accomplishes this feat.

Since people with problems in psychological functioning required help, whether or not this understanding had yet been reached, the gap was filled with best guesses; constructions of a 'psyche' based on non-falsifiable tenets (reification, see 3.2.1 and 3.3.2). In essence, the empty circle in figure 2, the one that depicts the brain, was filled with constructs thought to be able to perform this translation between input and output - the psyche -. These constructs were often based on subjective observations by psychotherapists. One such example would be Freud's (1923) construct of a psyche, based on ideas of 'layers in consciousness' he called the Id, Ego and Superego.

It seems that our knowledge about the brain has now started to reach a state where it may become possible to suggest and verify physical constructs capable of performing the same functions as hypothesised within the psyche constructs of various PT theories. In other words, it is becoming possible to fill the 'brain circle' with physically measurable structures and processes. Perhaps the time is right to translate the insights gained from brain research into physically defined constructs, which could then be used to further PT research, teaching and practice. Cozolino (2010) in his publication on neurology and psychotherapy has attempted to do just that.

The constructs as proposed by Cozolino (2010) can serve as an example of the type of constructs advocated for in this chapter. Even if the constructs he proposes turn out to be incomplete or inaccurate, they will still be a good example of what is advocated for here; since the fact that they *can* be proven wrong is a tribute to their usefulness. They can be improved upon, or be discarded, *because* they are physically measurable. To see how this is the case we shall discuss a description he provides on how a brain mediates between contextual input that is interpreted as threatening, and the resulting physical, behavioural and emotional output. This has been illustrated in figure 3.

According to Cozolino (2010) the sensory input received by the brain is first sent to the thalamus. There a quick, still unconscious, appraisal is made about the nature of the stimulus: threatening or non-threatening. If the stimulus could be considered threatening, two different pathways are stimulated: a fast and a slow one. The fast one leads directly to the amygdala, from where output is directly stimulated which can take the form of autonomic nervous system activation, behavioural responses and emotional responses. This fast circuit completely circumvents the consciously reasoning part of our brains. The slow circuit, however, sends the information on to the cerebral cortex. Here it can be more thoroughly, and thus more slowly and consciously, examined. Depending on the outcome of this secondary analysis, signals get sent on to the hippocampus and the amygdala to either inhibit or further stimulate the output already put into action.

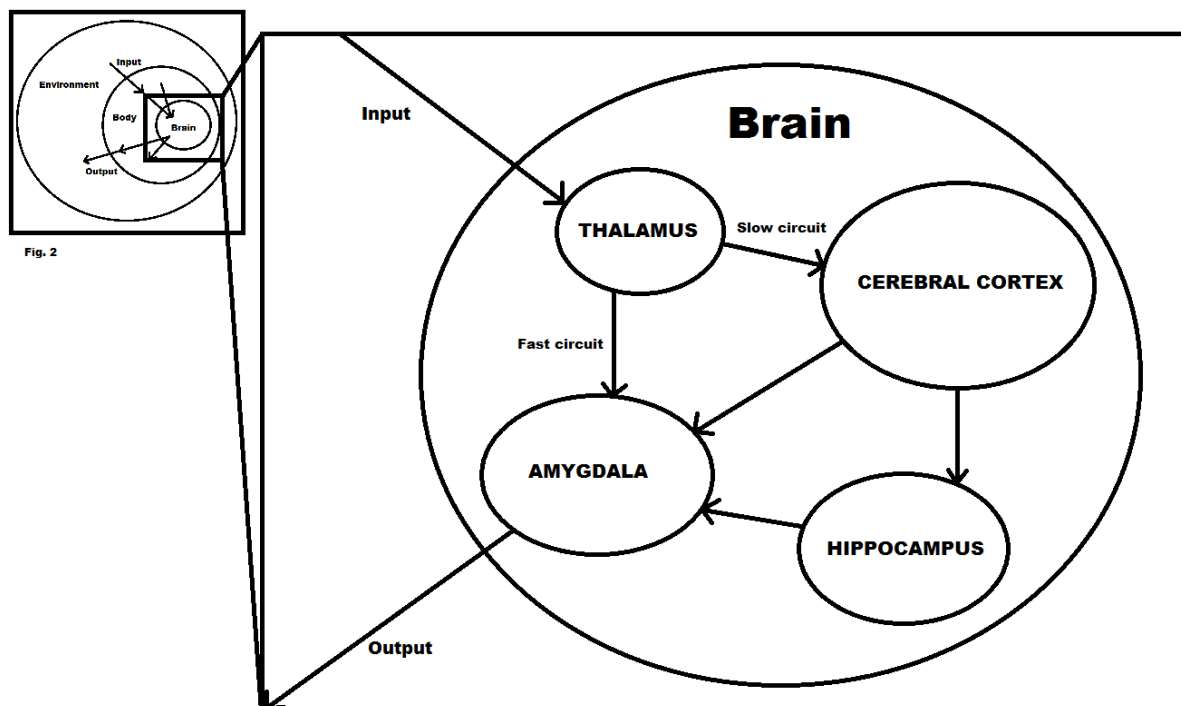


Figure 3: A simplified depiction of the process and brain structures involved in the perception and response to a potentially threatening situation (adapted from Cozolino, 2010, p. 243).

In theory we could measure someone's brain activity while they are subjected to a threatening situation. If the measured brain activity indeed follows the path as described in figure 3, then Cozolino's (2010) construct stands. If not, it can be discarded or amended and tested again. This is the beauty of such a definitively defined construct - it allows for definitive falsification. This in turn allows for the suggestion of falsifiable pathology theories and treatment mechanisms. If pathology, or the treatment targeting it, is supposed to alter something measurable within the circuit depicted in figure 3, then theories describing these too can be discarded or amended if they are falsified through brain imaging research. If this indeed manages to become reality, perhaps that would mean saying goodbye to the dodo-bird. Though, as we shall see in the next few paragraphs, this may not be quite as straightforward as that.

#### 4.1.2 Qualifying health

As discussed earlier (in paragraph 3.3.3) psychological functioning cannot objectively be judged 'good' or 'bad' without taking a perspective; good or bad for whom and when: the individual or the collective? Now or tomorrow? Presumably the systems currently in use within PT research, such as the DSM (American Psychiatric Association, 2015) and the



ICD/ICF (World Health Organisation, 2015), are also based on particular choices about perspective. Though this absence of an objective basis seems inherent to any definition of psychological functioning, to me it seems unwise to leave the subjective choices on which the definition rests implicit.

Whether the subjective choices on which our classification systems are based, were made consciously or not, is difficult to say. These choices and the reasons why they were made thusly are not explicated anywhere - as far as I have been able to find at least. As history has taught us though (see paragraph 3.3.3 for examples), these choices about pathology definitions matter.

There are for instance obvious ethical implications, because these definitions shape a key power dynamic in society; who is considered sick, who is supposed to do something about it, and even what they are supposed to be doing. Apart from this, however, implicit definitions are also ambiguous definitions. If two parties research the same phenomenon and use the same classification, but they are still able to have a different qualification of health in mind, this can lead to false research outcomes.

PT treatment research is intertwined with disease classification - the definition of a disease clearly stipulates the endpoint of a treatment after all. It is therefore my stance that classification systems should include an explication of the perspective underlying each disease construct. In this way researchers, but also policy makers, therapists and clients, would be aware of the chosen priority of interests for a particular treatment, and they could consciously choose whether the classification is right for their purposes. This could help in debating the merits of one definition over another - perhaps different definitions work better in different situations. However, more to the point for this thesis, it could help to get rid of theory ambiguity, and therefore to get rid of that pesky dodo.

#### **4.1.3 Alternative ways of looking at pathology and treatment theories**

Based on a falsifiable construct of the psyche, coupled with an explicitly stated concept of what is considered healthy functioning, pathology and treatment theories can be suggested. Cozolino's (2010) work, along with perhaps other similar endeavours, seems a good

candidate to base these theories on. Judging by own experiences in MPT education and from what I encounter in the literature though, these concepts have not yet replaced - perhaps are not ready yet to replace - the old psychopathology and treatment ideas of different PT schools. What seems to be standing in the way of replacing or merging these pathology and treatment theories into a more falsifiable and unified form?

Kradin (2008) argued that it may have something to do with the (exclusive) use of a reductionist paradigm in formulating most healthcare theories. He argued that most healthcare research is focussed on gaining understanding through reductionism; the idea that understanding of the whole can be gained by studying its components. Though he upholds that this is an important part of the scientific process, he argues that it cannot lead to full understanding of a complex - perhaps chaotic - system because it ignores important characteristics of such systems. It does not take into account the emergence of new system properties that arise from the totality of component interactions. For this he claims it is necessary to integrate other paradigms such as systems thinking and chaos theory into healthcare theory formulation.

Even though he does not explicitly state the same reasoning, Wampold (2001) seems to imply something similar to Kradin's (2008) standpoint when advocating for a paradigm shift from what he refers to as the 'medical model' towards what he styles the 'contextual model'. The medical model, he explains, is a way of explaining pathology and treatment effects based on linear interactions; pathology is in essence understood as being caused by a malfunctioning component, and a treatment is a specific intervention that fixes the component. The contextual model he advocates takes a more systemic viewpoint.

The brain is a complexly interacting system in which the different components influence each other's functioning. Pinpointing one specific cause of malfunctioning in such a system can be as difficult - and perhaps as meaningless - as pinpointing the start of a circular process. In such a complex system, something like the one illustrated in figure 2, it can make much more sense to look at the system as a whole and attempt to identify how the difficulties in functioning are maintained, and where the general opportunities can be located for guiding the system's functioning in a different direction. (Kradin, 2008; Cozolino, 2010).

Let us explore this idea through an example based on the depiction of psychological functioning in figure 2. The goal of healthcare in this example would be to improve the functioning of the whole system, i.e. the psychological functioning of an individual in his/her environment. As can be seen by the direction of the arrows in the illustration - though it is an oversimplified representation - this is a circular process: Input from environment > Body > Brain > Body > Output to the environment, etcetera. Any point in the chain can be the starting point for difficulties in the whole system. However, the same could be true for the reverse; any link in the chain may have the potential to positively affect the functioning of the whole.

Pathology and treatment ideas based on this way of thinking would not try to reduce understanding to a purely linear model, but rather attempt to identify generally which points in the system have the biggest potential negative effect, and at which points in the system relevant opportunities can be identified for instigating positive changes. Within the system of psychological functioning as illustrated earlier (figure 2), this could work as follows: pathology nosology could be based on a list of interacting variables that increase the chance of problems in functioning, such as specific irregularities in the brain, body and environment, or in the interaction between them (the input/output processes). Treatment mechanisms could be sorted and studied depending on the part of the system they target, and how well these interventions manage to affect the whole. For instance manipulating the content or intensity of chemical or sensory input, or through direct manipulation of the body or brain by for instance surgery.

Current theories seem to view PT as an intervention geared toward linearly fixing a faulty component of the psyche/brain. Perhaps instead PT should be viewed as an intervention in the client's environment (context) geared towards inducing positive change in the whole system's functioning. This in essence seems to be what Wampold (2001) is suggesting with his contextual model of therapy. PT can then be seen as just one of the many possible ways in which to influence psychological functioning. This system's based conception of treatment is illustrated in figure 4. It depicts PT as a form of sensory input manipulation, next to other methods as for instance medication to change body chemistry and surgery to alter brain physiology.

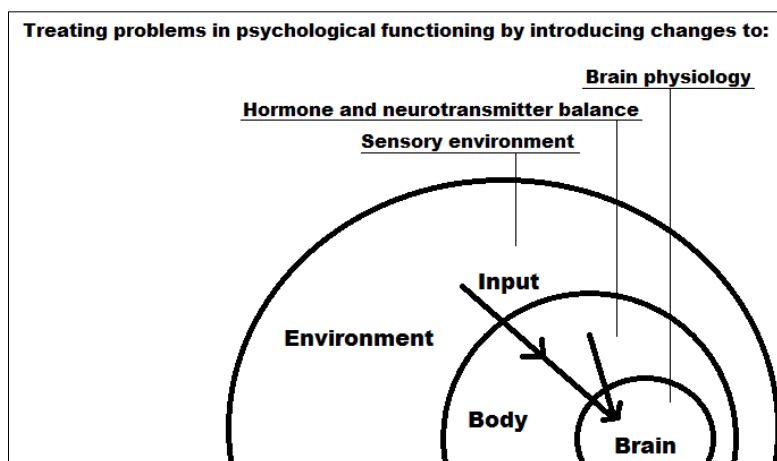


Figure 4: An illustration of potential treatment areas to be targeted for improving psychological functioning.

It is my hope that pathology and treatment theories which also take systems based reasoning into account will eventually replace or merge older theories. However, it is clear to me that we are not there yet. And though it may be tempting to discard old concepts all together, this may not be the best course of action either. For instance Wampold (2001) in his meta-analysis seems to find that the clarity of understanding gained by both therapist and client when following one specific PT theory - regardless which one -, seems to have an important positive effect on the outcome of the treatment. And Millon (2004) hastens to mention that even if the current pathology definitions are far from ideal, they do not completely lack merit. A clear grouping and naming of syndromes has great advantages in interdisciplinary communication and research, and could still lead to etiological insight and better treatments further down the road.

## 4.2 Improving experimental design?

The problems in PT research design that seem to have helped lead to the DBV (as discussed in paragraph 3.4), such as difficulties in constructing valid and reliable measurements, defining precise variables, and adequately controlling for confounding factors, can perhaps inspire a feeling of defeat. Maybe we simply do not (yet) have the capabilities to prove precise causal links between particular PT interventions and particular healthcare outcomes? As suggested in the previous paragraph, this may even inherently be impossible due to the chaotic nature of the psychological system.

However, the call for help with psychological difficulties does not vanish simply because there are no definitive solutions to these problems. In the following subparagraphs a few possibly helpful avenues will be discussed. Changing some elements within PT research - most notably how we choose to re-interpret research findings - can perhaps improve the situation.

#### **4.2.1 Exchanging linear certainty for system's based reasoning**

As is seemingly the case with weather patterns, the way human brains respond to and shape their environment, seems the outcome of such a complex (perhaps chaotic) system that making precise predictions about the effect of any one variable on the functioning of the whole is currently outside the realm of our capabilities (Kradin, 2008). Admitting uncertainty about results, however, should not mean giving up. When looked at from a different perspective, this simply opens up new avenues for pathology and treatment research.

According to Kradin (2008) the trick lies in looking at causality differently. Instead of trying to find precise causal links between independent and dependent variables, he suggests that our focus should shift towards measuring the *likelihood* of changing a dependent variable through the manipulation of an independent variable. In other words, only the type of hypotheses and the focus of experimentation need to shift.

Instead of looking for linear connections between variables, the goal of research becomes the identification of how system components and their interactions can contribute to the functioning or dysfunctioning of the whole system. It then becomes possible to identify potential entryways into the system and how interventions in these areas can have the biggest positive impact. This requires of both researchers and treatment providers that they let go of the false sense of certainty that our current therapy models seem to imply - that if a therapist does intervention A this will surely cause effect B.

In this changed framework intervention A has a certain *likelihood* of setting changes in motion in the whole system, which in turn may or may not cause effect B to occur. We could then study which factors in intervention A have the highest chance of producing effect B, and

refocus our efforts to understanding how and why this is perhaps different from one person to the next.

#### **4.2.2 A different perspective on placebos**

In the previous paragraph (4.1) the idea was introduced that we may have to change our perception of PT. We could start seeing it as a contextual treatment, as earlier depicted in figure 4. Herein our best knowledge about the effects of certain changes to sensory context could be used to affect a client's functioning. This would mean that PT would be conceptualised as a type of *contextual intervention*.

As I see it, other disciplines that also employ changes in a person's sensory environment to effect change could be seen as belonging to the same 'family' of interventions. Other types could include for instance: education - an intervention aimed at improving academic functioning through the manipulation of contextual variables -, mind-body medicine - aimed at improving physical functioning -, and marketing - aimed at steering consumer behaviour.

According to many researchers (see for instance: Moerman and Jonas, 2002; Kradin, 2008; Budd and Hughes, 2009) placebo effects, though often misunderstood and unfairly derided, can in essence be thought of as the effect of context on a test subject's perception. These changed perceptions can in turn - consciously or not - lead to measureable effects on physical or behavioural responses. For this reason Moerman and Jonas (2002) suggest using a more unambiguous name for these effects, suggesting the term 'meaning response'. Viewed as such, a contextual intervention is in a way the skilfully inducing of meaning responses to treat a perceived problem - regardless of being psychological, educational, medical or commercial in nature.

For some people this may seem a reason to discount meaning responses as a treatment avenue - perhaps they have feelings of unease about using a treatment associated with the placebo effect. There is, it seems to me, an aura of inherent deception associated with the term. I would posit that this association is due to a misunderstanding about the nature of meaning responses, and not something inherent to them.

According to de Craen Kaptchuk, Tijssen and Kleijnen (1999) historically the placebo response has been researched and studied mostly in the role of an annoying phenomenon that stood in the way of reaching valid conclusions about the benefit of medical treatments. In testing these treatments placebos would be (and still are) used to verify that any effect that is measured can be attributed to the medicinal substance under investigation, and not to meaning responses.

This type of verification process always requires deception. However, the way in which meaning responses are utilised in this situation, is about as far removed from how they are employed within a PT treatment, as is possible to be. The intended use in PT is one of treatment, not one of controlling an experiment. It is comparable to growing a flower; if you intended to grow potato plants in the same spot, the flower is considered a weed. If you wanted to grow exactly that flower - perhaps for its medicinal uses - it is something else entirely.

If ways can be found to make use of these meaning responses as a treatment vector without the use of deception, this would completely negate the uneasiness argument. Since the intended use is so very different from the setting in which the deception association was formed, I believe there is no reason to believe this could not be achieved. We should be looking to make the best possible use of this nifty healing response, and not be wasting energy deriding it because of ill-advised prejudice.

### **4.2.3 The role of context in healthcare research design**

Brain functioning can be affected by sensory context (Kradin, 2008). In return the brain can also affect the body's functioning (Benedetti et al., 2003). So, to be able to conclude which component of a treatment causes a certain effect on a conscious subject, the research must also take the subject's sensory context into account. This is why an experiment testing a treatment's efficacy - as for instance the RCT as it was discussed earlier (figure 1) - is always designed somewhat as illustrated in figure 5.

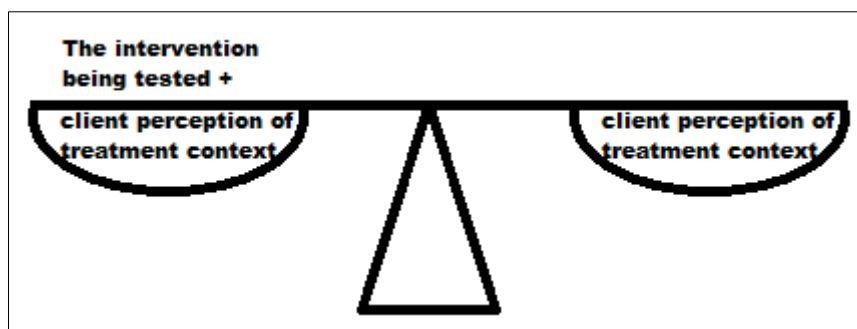


Figure 5: An illustration of a well-balanced healthcare research design.

According to Kradin (2008) in the situation as illustrated above any tipping of the scale can be ascribed to the intervention added on one side. This, however, is only the case if context was indeed controlled for properly - which is harder than one might think. If there is no difference, the efficacy must be sought in the natural healing process or in the context provided. He adds that it must be noted that this setup only allows for statements about the existence of an *added* benefit of the treatment. If the nature and relative sizes of the different effects need to be distinguished, a much more elaborate setup is needed, as for instance was done for the contextual and chemical effects of the administration of some anaesthetising drugs (Colloca & Benedetti, 2005).

As was determined earlier (paragraph 3.4.2 and 3.4.3), it is not yet possible to test the effects of naturalistic contextual variables within a PT treatment on psychological functioning. Even though very specific testing is not possible as of yet, there is no reason not to use a similar setup as discussed above to test the likely effects of contextual interventions on the whole system's functioning. The setup would then be similar to figure 6.

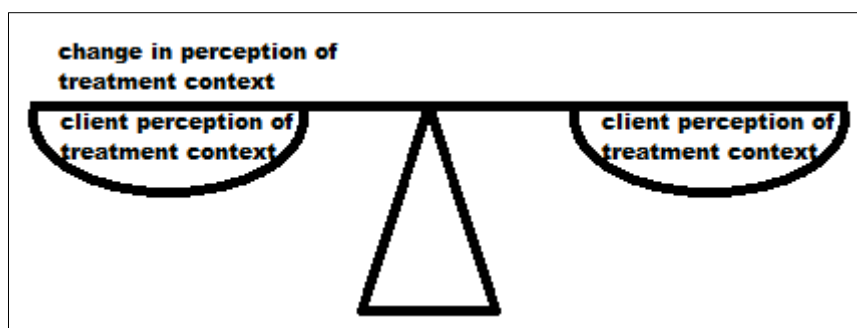


Figure 6: Testing the effect of different contextual components.



Other healthcare research already makes use of this setup, by for instance testing the effect of pill-colour on a medicine's effect (Blackwell, Bloomfield & Buncher, 1972). This particular research example involved inviting two different test groups to take pills with the same ingredients - but with a differently coloured exterior. If the groups are large and heterogeneous enough and the other circumstances under which the pills are taken do not vary much, then any significant difference must be ascribed to the participant's changed sensory experience.

New RCTs using these perspectives could be carried out to this effect, offering psychotherapists new data to base their treatment decisions on. A possibly limiting factor, however, could be ethical boundaries in healthcare research. Ethical limitations may prohibit the exclusion of factors that are suspected to be efficacious from treatments in clinical trials. There is however no limit on using older data to reach new conclusions. Insight may also be gained by examining and comparing research insights from other fields that also use contextual interventions. These were discussed previously and they included education, mind-body medicine and marketing.

Using older data to reach new conclusions is in fact what Wampold (2001) and Baskin et al. (2003) have been doing with their meta-analyses. By reanalysing older RCTs and other healthcare research, they were able to identify specific structural similarities between effective contextual treatments. This seemed to indicate that contextual factors such as certain aspects of the treatment's physical context, the therapists personal treatment style, alliance (a quantification of trust and agreement between therapist and client), and therapist allegiance (belief in efficacy) to their own treatment model, all seemed important factors in predicting treatment outcome.

## **5 BEYOND THE DODO-BIRD**

Our starting point involved an ethical and practical dilemma: whether, when the full significance of the DBV was taken into account, there would remain sufficient grounds (ethical as well as practical in nature) to warrant the continued prescription of (M)PT to clients. To resolve this issue the last two chapters (3 and 4) were dedicated to figuring out what the likely causes for the DBV in PT research were, and how the identified problems could perhaps be remedied. Having answered these two research questions, we can now attempt to answer the bigger question; to what extent does (M)PT stand as a valid option for treatment?

### **5.1 Moving towards a new framework for psychotherapy**

Kuhn (1996) offers a lens through which we can view progress in research. He suggests that, even though knowledge progression would ideally go in a linearly upward moving line, certain facets of human nature cause this process to falter. He implies that progress in human knowledge usually happens in incremental changes for a while, until suddenly great shifts occur. In a way the evidence that tries to nudge us towards the realisation that great changes are necessary, needs to build up a certain momentum before it is able to move us onto a new path. These big sudden changes are known as paradigm shifts.

Many of the sources discussed in this thesis (see for instance: Budd & Hughes, 2009; Cozolino, 2010; Frank, 1973; Frenkel, 2008; Kradin, 2008; Luborsky, Singer & Luborsky, 1975; Moerman & Jonas, 2002; Rosenzweig, 1936; Wampold, 2001) seem to agree that just such a shift is necessary in how we envision the inner workings of PT treatments. Throughout the preceding chapters, based on an amalgamation of their work, a new framework - perhaps even a new paradigm - for understanding PT has been emerging. By no means is any allusion made that this framework would form a comprehensive explanation of the intricate mechanics of a PT treatment. It may nonetheless be a solid enough - and more importantly a dodo-bird free - foundation to rebuild and mesh (M)PT theories on. It could therefore establish enough grounds for continued use and research of (M)PT.

A shift in conception was advocated, delineating a different way to understand how PT treatments function: shifting from a linear model, to a more system's based, circular model. In this framework a person's psychological functioning is visualised as a somewhat circular system comprised of environment, body, and brain (see also figure 2). This shift to a systems based framework also necessitates an adjustment in what we understand PT to entail. Herein PT can be understood to be the 'art' of changing a patient/client's experiences (context/environment) in order to induce changes in their psychological functioning (as depicted in figure 4).

As discussed earlier (4.2.1), however, in this new framework the surety that a certain change in a system will affect that system in a predictable fashion cannot be ascertained - not yet, and perhaps never. Given this uncertainty at the core of our new foundation, it may seem that the question of whether or not (M)PT remains a valid treatment option, must be met with a negation. To prevent a premature rejection of this treatment avenue, however, this paragraph will examine whether this uncertainty can be ameliorated. In an attempt to mitigate this vulnerability, psychology research and evolutionary reasoning will be used to see if this weak spot can be fortified sufficiently.

### **5.1.1 Contextual components in human brain development**

The rather imprecise term 'art' in the above described conception of PT was chosen intentionally. It accurately depicts the strangely uncertain task faced by therapists in this new framework; to choose an intervention in the client's environment without the certainty that intervention A will always lead to result B in that person's psychological functioning.

This uncertainty seems a formidable obstacle to the reliable implementation of contextual treatments such as PT. There may be a way to reduce that obstacle in size. It was argued earlier (paragraph 4.2.1) that in theory it would be possible to identify components within a system that have a higher likelihood, when manipulated, to cause a certain result. In other words, it may be possible to figure out which contextual components in a person's system of psychological functioning (as depicted in figure 2) have the highest chance of causing positive changes when manipulated as a therapeutic intervention.

Knowing this would make a therapist's job much more achievable. A client's environment is built from a million different components - it would help to know whether a therapist for instance needs to worry as much about the colour of the curtains, as about body language and words. The brain imaging technology needed to accurately make this distinction, however, is not around yet (as discussed in paragraph 3.4.2). We are therefore, once again, forced to find a less direct route to figure out which components of a person's environment have the highest chance of affecting someone's psychological functioning in a positive way. We shall therefore take the route of Darwinian logic, coupled with outcomes of biology and psychology research, to try making this distinction from basic premises.

Since Darwin's days much has been discovered about the mechanism behind the observed changes between and within species. For instance we now know about genetics: that each living cell contains a heritable code (DNA) that to a high degree determines what the whole organism looks like and how it behaves. However according to Waddington (2012) the conversion of DNA code - in essence just information - into a complete organism, is not quite so straightforward. Simply 'reading' the instructions within the code - the genotype - does not provide a complete map of all the organism's physical and behavioural traits - the phenotype. Waddington suggests that a whole host of developmental processes lie between the genotype and the eventual form of the phenotype. He suggests the term 'epigenotype' for the description of these processes.

Waddington (2012) explained that the epigenotype for some parts of our bodies, is not only a conversion of DNA instructions into physical structures, but is the product of an interaction between DNA instructions and the *environment*. Put differently, a part of the translation process between DNA and eventual physical appearance and behaviour, seems determined by the experiences we have. According to van Elk and Hunnius (2010) and Swaab (2010) this is particularly true for human brains. Van Elk and Hunnius (2010) explained that babies are born with relatively underdeveloped brains, which then are moulded into their eventual shape through experiences. This may sound like a disadvantage, but it in fact seems to be what allows our species to adapt so easily to new circumstances.

The picture of human brain development sketched by van Elk and Hunnius (2010), along with many other sources on brain development (for instance: Swaab, 2010; Cozolino, 2010)

suggests that experiences are, in a way, to human brains what water is to a river. Sensory input as it were ‘flows’ into the brain, follows a meandering path through the brain and flows on in the form of activation or inhibition of other organs or muscles. However, these experiences do not flow passively through the brain, just as water does not flow *through* a river. Like a river in a way *is* the flowing water, one could say that a brain *is* its experiences. Just as water deposits sand and erodes the riverbanks as it flows, shaping and changing the river as it goes, so do our experiences literally shape and grow our brains.

Not only does this happen in baby-brains, it continues to happen, to a lesser extent, in adults throughout their lifetimes (Swaab, 2010). This is what allows new experiences to (re)shape our brains and therefore our behavioural patterns. According to Cozolino (2010) when a psychotherapist offers a client new experiences, this is how the changes are caused.

### **5.1.2 The importance of human interaction**

Seeing experiences in this new light, allows us to wonder which types of experiences are most important for the development of healthy human brains with corresponding healthy behaviours. From an evolutionary perspective, a brain can be seen as an organ that evolved to aid organisms in their survival in particular environments. Having a brain allows them to respond to differences in the environment by interpreting sensory information, and by instigating reactive behaviours based thereon. Brain evolution of a particular species will therefore be an adaptation to the particular environment in which that species has survived over the years: a bird’s brains will have become good at interpreting and responding to differences in the air, just as a fish’s brains will have adapted to do something similar in the water.

According to Hood (2014) people have evolved to navigate complex social environments, similar to how birds evolved to navigate the air. The type of behaviour we logically would associate with a well functioning human adult is therefore adequate social behaviour. This is the type of behaviour logically targeted by PT: internal (e.g. thoughts or feelings) or external behaviour that hinders a person’s own or their social environment’s functioning (e.g. criminal, intrusive or disruptive actions).

To understand how these behaviours are developed in human beings during their lifetimes, we can turn to developmental psychologists, who have attempted to answer just this question. Many different experiments have shown that for social mammals such as apes - humans included -, healthy brain development seems most dependent on a particular type of experience (for instance: Harlow, 1958; Ainsworth, 1978; Bachevalier, Alvarado & Malkova, 1999; Kaufman, Plotsky, Nemeroff & Charney, 2000; Chugania, Behena, Muzika, Juhásza, Nagya & Chugania, 2001).

The type of experience referred to here, is the type that promotes what psychologists call 'attachment': Being touched, held, talked to, etcetera. At the beginning of this chapter the question was posed whether a hierarchy can be said to exist in the importance of different types of contextual components when trying to achieve changes in psychological functioning. According to developmental psychology the answer thus seems to be that interaction between people trumps curtain colour.

Based on similar reasoning Cozolino (2010) conceptualises PT as involving the 'social synapse' in a treatment. His concept suggests a parallel between people and neurons, in that they both function as part of a network (system), and that much of the network's functioning happens in the space between the neurons/people. A neuron reacts to stimulation in the form of neurotransmitters (signalling chemicals) released by its neighbouring neuron into the synapse between them. Future activation of the neuron partly depends on the history of interactions with its neighbours. Cozolino suggests that people operate in a somewhat similar manner. Much of their behaviours are the consequence of interactions with the people around them, in addition interactional patterns that endure have a lasting effect in the form of recurring behavioural patterns - learnt social, cognitive and emotional behaviours.

How and why an attachment promoting setting manages to generate healthy brain development is a whole different thesis. However according to Cozolino (2010) it is within this type of setting (even as adults) that we best learn new adaptive social behaviours - both the internal ones (e.g. thoughts, emotional responses) as well as the external ones (e.g. visible/audible behaviour). This is corroborated by Wampold's (2001) meta analysis, where it was concluded that these very basic components of attachment building interaction, such as

finding agreement and communicating safety, seem to be among the most important components determining PT efficacy.

### **5.1.3 Psychotherapy: a valid treatment option?**

The preceding chapters and paragraphs have sketched a picture of the current state of PT theory and research. An analysis was given of what could have led to the DBV predicament in the first place and which steps could be taken to get around the issue. A new framework for understanding PT emerged, but one with uncertainty at its core. The present chapter then tried to see to what extent this uncertainty could be ameliorated. This leaves us to now consider an answer to the question at the heart of this thesis; is a shift towards the new framework, and are the arguments presented to mitigate the uncertainty at its heart, enough to continue prescribing PT conscientiously?

Let us start with a point about contextual treatments in general. It may not have been stated explicitly enough, but the added benefit of this type of treatment seems actually to not be under fire at all. Every research report that I could find on the outcome differences between a contextual treatment being offered and nothing being offered - even if it just involved the experience of being prescribed, and taking, a placebo pill - seemed to show that the contextual treatment offered important benefits over the act of offering no help at all. See for instance the work of Moerman and Jonas (2002) for a summation. This makes so much common sense, that we can easily forget to be amazed by it.

A different way to look at it is that the DBV did not affect contextual treatments in general, just the distinction between the supposed different types of these treatments. The real question in essence became whether it can be shown that PT, in one form or another, can actually *beat* the simple setting of the ‘pill prescribing doctor’ in effectively harnessing contextual components to treat problems in psychological functioning. I would posit, based on the studies read for writing this thesis, that perhaps research based on the old paradigm was - for the most part - not even aimed at finding out how to best utilize contextual components in PT. Outcomes of these studies are therefore not useful to determine whether they can be, or already are, more effective.

This thesis showed that treating problems in psychological functioning through contextual interventions is not straightforward. Through exploring basic psychology and biology research, however, some tentative suggestions could be made to mitigate this uncertainty somewhat. We can seemingly infer a few things: that brain development thrives best in an attachment promoting settings, and that a treatment process needs to be dynamic and at least somewhat tailored to the individual. This last point because we now understand that brains *are* in a way the person's experiences. Since everyone lives their life through their own set of unique experiences, therapy cannot be a catch all intervention.

If the new PT research aim were to become a quest to design the most effective way(s) to use contextual interventions to positively influence a client's psychological functioning, then a good starting point might be the wealth of information - the rich descriptions of therapist and client experiences - in the existing PT literature. Most, if not all, PT theories that formed the older framework seem to fit quite nicely into the new one as well - though they may require some reinterpretation. Whether their specific interventions end up being supported by empirical evidence, however, remains to be seen.

Absolute certainty cannot yet be reached about the value of PT as a treatment option. Our technological capabilities are not yet sophisticated enough to test hypotheses sensitively enough. I hope, however, that it has been shown that PT offers something substantial and real. Furthermore, the benefits of having a treatment option for psychological anguish that does not involve chemical or physical alterations of a person seem to far outweigh the risks of continuing the exploration of this treatment avenue. This seems a reasonably strong enough fortification for us to continue prescribing and researching PT.

## **5.2 Does music psychotherapy have a place within the new framework?**

Now that we are nearing the end of this thesis, the time has come to make our way back to the music, since this thesis was written as part of a music therapy Master's degree. Let us for a moment assume that the case has been made for psychotherapy to remain a valid treatment option. How then does this validation translate to *music* psychotherapy? Can music occupy a sensible place in the new PT framework? In an earlier paragraph (5.1.2) a model for understanding the role of human interaction within therapy was discussed; the idea of a social



synapse (Cozolino, 2010). I propose to base understanding of the role of music in therapy on this concept.

The contribution of music to PT as I propose it necessitates a discussion of the concept of a 'medium' - the physics term, not the supernatural kind. Take for instance the way that information can travel from one place to the next through sound; it requires something to travel 'through'; a medium. Speech is a (deliberately) manufactured disturbance that propagates - often through air, the medium - before it reaches someone's ears, where it can subsequently be decoded for the information it may carry.

Cozolino's (2010) idea of the social synapse also seems to require a medium. Just as the synapse between neurons needs to be bridged for information to be passed on, so must the 'synapse' (gap) between people's brains be bridged for communication to happen. What does that information travel through? For neurons it seems to require neurotransmitters, this would be their bridging medium. Between people perhaps we can picture the medium as the sensory perceptions we pick up from the actions of people around us, such as visual, tactile, auditory, or even olfactory cues. Let us call this the sensory medium between people. I imagine that this medium communicates the intentions and internal representations from one person to another, because both people are capable of imagining what they would be feeling if their own face or body moved in the way the other person's did.

Although PT schools have traditionally focussed more on sending verbal/lingual cues through the sensory medium to induce changes in clients, using words is not the only way in which people communicate meaning between each other. For instance when we are touched softly, or spoken to in a soft voice, we can interpret the other person's intent as 'caring'. Since these types of cues seem just as capable of fostering an attachment promoting environment, it follows that other forms of communicative cues could also be exploited to perform a PT treatment with.

According to Smeijsters (2006), aside from being able to directly elicit physiological reactions and structure mental and physical processes, music is able to illicit memories, values and associations, and be used as symbol, metaphor and analogy. This potentially makes music a good contender for use in PT treatments, as music therapists have (intuitively) recognised

already. Aside from music though, this may also be true for many other creative media forms, such as visual arts, drama, and body movement - dance, play, sports, touch just to name a few.

One could wonder - and perhaps this should be studied further - whether for some people, or under certain circumstances, making use of these other media to induce the desired meaning responses could be more effective than being limited to verbal interventions. Though the traditional models of PT may have mostly kept their treatment investigations confined to the use of verbal/lingual interventions, creative therapists (music therapists included) have done much pioneering work in this area already. Whether or not MT research can sufficiently show this benefit, again remains to be seen. However, in theory, based on the reasoning provided here, it seems plausible that it could provide an added benefit. The use of music should therefore not be discarded out of hand.

## 6 SUMMARY AND REFLECTION

When various forms of psychotherapy (PT) are scientifically compared in terms of efficacy, no clear ‘winner’ emerges. This is known as the dodo-bird verdict (DBV), named after the dodo-bird in Lewis Carol’s (1865) *Alice in Wonderland*, who awards prizes to everyone in a race with no clear winners. This verdict casts serious doubts on the credibility of PT as a treatment. It raises both ethical and economic issues; why would anyone choose, let alone pay for, a treatment that does not stand up to scientific scrutiny? Since many forms of music therapy make use of psychotherapy techniques and theories, this is a problem that also affects music therapists. Understanding *what could have caused the DBV to occur*, and *finding solutions to the situations that seem to have caused it*, are essential for the viability of the field.

### 6.1 Research methodology

When many researchers try to answer (parts of) such a complex question, this can lead to a situation where one cannot see the forest for the trees. The dodo-bird verdict seems the outcome of just such a situation. The sheer amount of data could be causing therapists and researchers to miss the bigger picture. In this thesis conceptual research methods were used to attempt forming this bigger picture; psychotherapy theory and research were explored to find causes of and solutions to the DBV situation.

Though this research process uncovered many critiques on PT theories and research, it also found much support for the efficacy of PT treatments nonetheless. Primarily this research showed a need to thoroughly re-conceptualise how we understand therapy to work and how we can best research it, recognising it as a contextual treatment of the brain/social system.

First possible causes were examined, for which a broader view was taken on the DBV. In this thesis it is seen as the confirmation of a null-hypothesis where common sense understanding of the situation would overwhelmingly suggest this should not be the case. Any research outcome where this applies could therefore be classified a dodo-bird verdict. This led to exploring the philosophical underpinnings of empirical research in general, and applied science in particular, to identify possible weak spots in theory and/or research that could

hypothetically lead to a DBV. Armed with this understanding the specific situation in PT theory and research were examined to see which of the hypothetical weak spots applied, and what could be done to remedy the situation.

## **6.2 Examining foundations**

### **6.2.1 Identifying weaknesses in theory construction**

The weaknesses identified in PT theory concerned difficulties in defining pathology and health. Scientists have already long inferred that the origin of psychological problems should likely be sought in brain functioning. The problem with this has been that researchers had not yet been able to measure brain functioning accurately enough to base psychopathology theories on falsifiable descriptions of these processes.

To bridge the gap between what we did not know about the brain, and the need that existed for treatment, PT theoreticians had to suggest *non-falsifiable* constructs that provided explanations for what was causing the symptoms. This in turn may have, at least in part, led to the situation we face today; many different PT theories that cannot be disproven because of the way they define the psyche - as something not measurable physically, with matching non-falsifiable treatment mechanisms. In addition to that, the definition of what is considered a healthy state is, as the word 'considered' suggests, dependent on the perspective one takes. This can lead to ambiguities in how outcomes are interpreted when this is not explicitly reported or considered, which may have contributed to the occurrence of a DBV as well.

To address these issues psychopathology needs to be based on a falsifiable construct of the psyche. This would involve using descriptions of physically measurable brain structures in explanations of pathology. This allows for theories about psychological functioning and treatments to be falsified based on brain research, thereby making it possible to eliminate theories by testing them. Furthermore psychological health should be defined with an *explicit* statement of perspective, making it possible to debate the merits of choosing certain interests over others. This explicit statement also enables us to measure and discuss treatment and research outcomes with less ambiguity. In the transition towards using falsifiable psyche

definitions steps have been taken. However, technical imaging of brain processes still limits what can be achieved.

### **6.2.2 Identifying weaknesses in psychotherapy research**

Besides weaknesses in theory construction, the effect of weaknesses in research implementation on the occurrence of a DBV were also considered. Problems were identified in four elements: performing objective measurements, defining useable variables, ethical constraints, and controlling for confounding factors.

Measuring psychological functioning validly and reliably is challenging. The lack of falsifiable definitions for the psyche and psychopathology, lead to a whole host of problems in measuring them. It is for instance rather difficult to measure something that is arbitrarily defined on observations of symptoms, as for instance the American Psychiatric Association's diagnostics and statistical manual (DSM) is, instead of physically defined on the basis of brain functioning.

Additionally the fundamental difficulty of designing valid and reliable measurements of something that is fundamentally not verifiable - such as a person's conscious experiences and thought processes - raises even more obstacles. Overcoming this fundamental difficulty completely does not seem possible, the best way forward however seems to be research that correlates reported experiences with brain imaging to make better models of psychological functioning. Though advances are being made in this area, our technology is not yet at a level that this can be done well in naturalistic settings.

The variables used in PT research present their own set of challenges. In current PT research diagnoses, such as for instance the DSM categories, are often used as dependent variable. This is problematic because these diagnoses are syndromal groupings arranged not by logical etiological principals, but by arbitrarily grouping symptoms together. This leads to high comorbidity and inconclusive diagnoses, making these diagnoses unsuitable as dependent variable. This makes it likely that outcome measures do not accurately represent the outcome of treatment.

A similar problem occurs when using an entire PT treatment as independent variable. A PT treatment usually takes several sessions over an extended period of time, and within one research trial more than one therapist is often used to deliver the treatment. This introduces so many possibilities for differences between participants that it becomes impossible to make confident assertions about causality.

Another hurdle that also decidedly complicates neutralising the DBV situation are ethical restrictions in PT research. Though the societal necessity of ethical restrictions in healthcare research is obvious, it nonetheless forms a formidable obstacle for devising valid and reliable experiments. Looking at old research data with new analysis tools may be a partial way around this, but fundamentally this means that some things cannot easily - perhaps should not - be tested.

Differences between experimental conditions that can potentially affect the outcome of an experiment are known as confounding factors. One such factor is perhaps being underappreciated or misunderstood in many PT research trials: contextual factors. Neurological studies suggest that brain functioning cannot only be influenced through the introduction of chemical compounds or surgery, but also through introducing changes to the brain's sensory context. Much - perhaps all - of PT efficacy is likely attributable to this mechanism of change. Re-imagining PT as a contextual treatment and setting up research and (re)writing theories accordingly could go a long way in solving the DBV situation, and thereby validating PT as a viable treatment option.

The problem remains that making accurate statements about causality within healthcare is difficult - if not impossible - without overcoming the inherent difficulties in devising falsifiable psyche and psychopathology constructs, research measurements, variables and controls. However, perhaps progress can be made without perfect measurements and variables. Since linear modelling seems not to lead to full understanding of psychological functioning, we could consider shifting to other research paradigms to break this impasse in the field, such as systems thinking and chaos theory.

In this scenario many research methods can remain largely the same. The interpretation of data, however, would change significantly. Instead of searching for direct causal links

between tightly defined variables, one could try to understand the whole system. Instead of trying to find one specific component in the equation to change, an overview could be made of different components and their potential to affect the whole system's functioning.

Though a few weaknesses remain, the final verdict on whether or not PT can be conscientiously prescribed, is a positive one. If the switch is made to the herein proposed new framework for understanding the inner workings of PT, then the dodo-bird problem seems to disappear and many new research avenues open up.

### **6.3 The music in music psychotherapy**

The specific role of music in therapy was not the focus of research. During the research process, however, a new perspective on the role of music in therapy was reached, motivating the addition of a final paragraph dedicated to this topic. Within the theoretical framework of PT as a contextual treatment, music could be viewed as an alternative or complementary medium through which the treatment can take place.

PT treatments - MPT included - generally aim to change internal (e.g. thoughts and emotions) and external behaviour and experience that negatively impacts a client's and/or their social system's functioning. Developmental psychology research seems to show that the most effective way to impact this type of functioning, is through attachment strengthening interactions with another person. These interactions take place through a sensory medium, in the form of for instance visual, audio or tactile stimulation.

Researching and harnessing more from all types of human communication modes could offer PT practitioners better - and more - tools for their trade. Related fields such as music or arts therapies, education and advertising may have much to bring to the table in this exploration, since they have already been explicitly studying and using alternate communication modes to reach their target audiences for a longer time.

## 6.4 Reflection

Reflecting back on the process of writing this thesis prompts me to add a few final thoughts in conclusion. This process has been a kind of personal journey for me. I needed to find answers to the questions posed in this thesis before I could reconcile my chosen profession with my wish to help people; the dodo-bird verdict in my perception cast doubt on the ethics of prescribing any form of psychotherapy as a treatment. I feel my understanding has grown and changed significantly over the past few years, and I feel that - at least for now - I have found satisfactory answers. I am, however, aware that this thesis is by no means a final, nor a comprehensive answer to the questions posed, and I therefore do not consider this the journey's end.

There were many theoretical components that deserved a much more in depth analysis and inclusion in this thesis, the common factors model, chaos theory and systems thinking are a few of these subjects. Many more sources and research areas could probably have been included in the analysis as well, but time and resource constraints limited the thesis' scope. It happened many times during the process that a new piece of information turned my understanding of the puzzle upside down, which meant I had to sit down and rearrange all the pieces in my head once more. I have no reason to believe that now - the moment I choose to finish the thesis - is somehow different in that I cannot thoroughly change my mind again.

It is my hope that whomever ends up reading this thesis, is at least shaken in their certainty about what they know - or think they know - about psychotherapy. I have come to believe that this kind of 'earthquake' in the assumptions we hold dear, is the only way forwards to better understanding and better treatments.



## REFERENCES

- Aigen, K. (2012). Reading 2: Philosophical inquiry. In K. E. Bruscia (Ed.), *Readings on music therapy theory* (pp. 34-49). Gilsum, NH: Barcelona publishers.
- Ainsworth, M. D. (1978). The bowlby-ainsworth attachment theory. *Behavioral and Brain Sciences*, *1*(3), 436-438.
- American Psychiatric Association. (2014). DSM. Retrieved: 29<sup>th</sup> December 2014. Retrieved from: <http://www.psychiatry.org/practice/dsm>
- Bachevalier, J., Alvarado, M. C., & Malkova, L. (1999). Memory and socioemotional behavior in monkeys after hippocampal damage incurred in infancy or in adulthood. *Biological Psychiatry*, *46*(3), 329-339. DOI: 10.1016/S0006-3223(99)00123-7
- Baskin, T. W., Tierney, S. C., Minami, T., & Wampold, B. E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology*, *71*(6), 973-979. DOI:10.1037/0022-006X.71.6.973
- Beauchamp, T. L. (2001). *Philosophical ethics: An introduction to moral philosophy* (3rd ed.). New York, NY: McGraw-Hill.
- Benedetti, F., Maggi, G., Lopiano, L., Lanotte, M., Rainero, I., Vighetti, S., & Pollo, A. (2003). Open versus hidden medical treatments: The patient's knowledge about a therapy affects the therapy outcome. *Prevention and Treatment*, *6*, Article 1.
- Berlin, H. A. (2011). The neural basis of the dynamic unconscious. *Neuropsychoanalysis*, *13*(1), 5-31.
- Blackwell, B., Bloomfield, S. S., & Buncher, C. R. (1972). Demonstration to medical students of placebo responses and non-drug factors. *The Lancet*, *299*(7763), 1279-1282.
- Bruscia, K. E. (1998). *The dynamics of music psychotherapy*. Gilsum, NH: Barcelona Publishers.
- Bruscia, K. E. (2012<sup>b</sup>). Reading 1: Developing theory. In K. E. Bruscia (Ed.), *Readings on music therapy theory* (pp. 20-33). Gilsum, NH: Barcelona publishers.
- Bruscia, K. E. (2012<sup>a</sup>). Introduction. In K. E. Bruscia (Ed.), *Readings on music therapy theory* (pp. 15-18). Gilsum, NH: Barcelona publishers.
- Budd, R., & Hughes, I. (2009). The dodo bird verdict - controversial, inevitable and important: A commentary on 30 years of meta-analyses. *Clinical Psychology and Psychotherapy*, *16*, 510-522.

- Carroll, L. (1865). A caucus-race and a long tale. In: Carroll, L., & Tenniel, J. (2003). *Alice's adventures in wonderland, and through the looking glass*. New York, NY: Barnes & Noble Classics. (Original work published 1865)
- Chabris, C., & Simons, D. (2009). *The invisible gorilla: How our intuitions deceive us*. New York, NY: Broadway Paperbacks.
- Chugania, H. T., Behena, M. E., Muzika, O., Juhásza, C., Nagya, F., & Chugania, D. C. (2001). Local brain functional activity following early deprivation: A study of postinstitutionalized Romanian orphans. *Neuroimage*, *14*(6), 1290-1301. DOI: 10.1006/nimg.2001.0917
- Colloca, L., & Benedetti, F. (2005). Placebos and painkillers: Is mind as real as matter? *Nature Reviews Neuroscience*, *6*, 545-552.
- Cozolino, L. (2010). *The neuroscience of psychotherapy; healing the social brain* (second edition ed.). New York, NY: W. W. Norton & Company.
- Darwin, C. R. (1869). *On the origin of species by means of natural selection, or the reservation of favoured races in the struggle for life*. (5th ed.). London: John Murray.
- de Craen, A. J. M., Kaptchuk, T. J., Tijssen, J. G. P., & Kleijnen, J. (1999). Placebos and placebo effects in medicine: Historical overview. *Journal of the Royal Society of Medicine*, *92*, 511-515.
- Edwards, J. (2002). Using the evidence based medicine framework to support music therapy posts in health care settings. *British Journal of Music Therapy*, *16*(1), 29-34.
- Fachner, J., & Stegemann, T. (2013). Electroencephalography and music therapy: On the same wavelength? *Music and Medicine*, *5*(4), 217-222. DOI: 10.1177/1943862113495062.
- Frank, J. D. (1973). *Persuasion and healing: A comparative study of psychotherapy* (Revised edition). Baltimore, MD: Johns Hopkins University Press.
- Freeman, W. J. (2004). How and why brains create meaning from sensory information. *International Journal of Bifurcation and Chaos*, *14*(02), 515-530.
- Frenkel, O. (2008). A phenomenology of the 'placebo effect': Taking meaning from the mind to the body. *Journal of Medicine and Philosophy*, *33*(1), 58-79.
- Freud, S. (1923). Das ich und das es. *Internationaler Psychoanalytischer Verlag*, Leipzig, Vienna and Zürich.

- Freud, S. (1968). Project for a scientific psychology. In J. Strachey (Ed.), *New introductory lectures on psychoanalysis: Standard edition of the complete psychological works of Sigmund Freud* (vol. 22). (pp. 3-182). London: Hogarth Press.
- Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: An introduction* (2nd ed.). New York, NY: SAGE Publications.
- Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *British Medical Journal*, 331(7524), 1064-1065.
- Harlow, H. F. (1958). The nature of love. *American Psychologist*, 13(12), 673-685.
- Hood, B. (2014). *The domesticated brain; A pelican introduction*. London: Pelican books.
- Kahneman, D. (2012). *Thinking fast and slow*. London: Penguin books.
- Kaplan, R., & Saccuzzo, D. (2012). *Psychological testing: Principles, applications, and issues* (8th ed.). Boston, MA: Cengage Learning.
- Kaufman, J., Plotsky, P. M., Nemeroff, C. B., & Charney, D. S. (2000). Effects of early adverse experiences on brain structure and function: Clinical implications. *Biological Psychiatry*, 48(8), 778-790. DOI: 10.1016/S0006-3223(00)00998-7
- Kradin, R. (2008). *The placebo response and the power of unconscious healing*. New York, NY: Routledge.
- Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: University of Chicago Press.
- Kutchins, H., & Kirk, S. A. (2003). *Making us crazy: DSM: The psychiatric bible and the creation of mental disorders* (3rd ed.). New York, NY: Free Press.
- Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes"? *Archives of General Psychiatry*, (8), 995.
- Millon, T. (2004). *Masters of the mind; exploring the story of mental illness from ancient times to the new millennium*. Hoboken, NJ: John Wiley & Sons.
- Millon, T., & Simonsen, E. (2010). A précis of psychopathological history. In T. Millon, R. F. Krueger & E. Simonsen (Eds.), *Contemporary directions in psychopathology; scientific foundations of the DSM-V and ICD-11* (pp. 3-52). New York, NY: The Guilford Press.
- Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding the meaning response. *Annals of Internal Medicine*, 136(6), 471-476.

- Otera, M. (2013). Is the movement of evidence-based practice a real threat to music therapy? *Voices: A World Forum for Music Therapy*, 13(2)
- Oxford dictionary (2014). Definition of knowledge. Retrieved, 30-06-2014, Retrieved from: <http://www.oxforddictionaries.com/definition/english/knowledge>.
- Popper, K. R. (1972). *Objective knowledge; an evolutionary approach*. Oxford: Oxford University Press.
- Rescher, N. (2003). *Epistemology: An introduction to the theory of knowledge*. Albany, NY: State University of New York.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6(3), 412-415.
- Schneider, D., & Lilienfeld, D. E. (2015). *Lilienfeld's foundations of epidemiology*. (4th ed.). New York, NY: Oxford University Press.
- Shapiro, D. A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin*, 92(3), 581-604.
- Smeijsters, H. (2006). Verklaringsmodellen. In H. Smeijsters (Ed.), *Handboek muziektherapie; evidence based practice voor de behandeling van psychische stoornissen, problemen en beperkingen*. (pp. 90-114). Houten: Bohn Stafleu van Loghum.
- Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752-760.
- Smolik, P. (1999). Validity of nosological classification. *Dialogues in Clinical Neuroscience*, 1(3), 185-190.
- Swaab, D. (2010). *Wij zijn ons brein; van baarmoeder tot alzheimer*. Amsterdam: Uitgeverij Contact.
- Thyer, B. A. (2001<sup>a</sup>). Introductory principles of social work research. In Bruce A. Thyer (Ed.), *The handbook of social work research methods* (pp. 2-25). Thousand Oaks, CA: SAGE Publications.
- Thyer, B. A. (Ed.) (2001<sup>b</sup>). *The handbook of social work research methods*. Thousand Oaks, CA: SAGE Publications.
- Thyer, B. A., Epstein, W. M., Sowers, K., Ellis, R. A., Meyer-Adams, N., Graham, J. R., Al-Krenawi, A.; Turner, F. J. (2001). Part III conceptual research. In Bruce A. Thyer (Ed.), *The handbook of social work research methods* (pp. 366-425). Thousand Oaks, CA: SAGE Publications.

- Trochim, W. M. K. (2006<sup>a</sup>). Research methods knowledge base; reliability. Retrieved from: <http://www.socialresearchmethods.net/kb/reliable.php>
- Trochim, W. M. K. (2006<sup>b</sup>). Research methods knowledge base; measurement validity types. Retrieved from: <http://www.socialresearchmethods.net/kb/measval.php>
- van Elk, M., & Hunnius, S. (2010). *Het babybrein; over de ontwikkeling van de hersenen bij baby's*. Amsterdam: Uitgeverij Bert Bakker.
- Vink, A., & Bruinsma, M. (2003). Evidence based music therapy. *MusicTherapy Today*, 4(4), 1-26.
- Waddington, C. H. (2012). The epigenotype. *International Journal of Epidemiology*, 41(1), 10-13. Reprint - first published in 1942
- Waldman, G. (2002). *Introduction to light: The physics of light, vision and color*. (2nd ed.). New York, NY: Dover publications.
- Wampold, B. E. (2001). *The great psychotherapy debate; models, methods and findings*. New York, NY: Lawrence Erlbaum Associates.
- Wigram, T. (2014). The religion of evidence-based practice: Helpful or harmful to music therapy? In J. de Backer, & J. Sutton (Eds.), *Music in music therapy : Psychodynamic music therapy in Europe: Clinical, theoretical and research approaches* (pp. 238-259). London: Jessica Kingsley Publishers.
- Wigram, T., Pedersen, I. N., & Bonde, L. O. (2002). *Comprehensive guide to music therapy : Theory, clinical practice, research and training*. Philadelphia, PA: Jessica Kingsley Publishers.
- World Health Organisation. (2015). Classifications. Retrieved 21<sup>st</sup> of November 2015. Retrieved from: <http://www.who.int/classifications/en/>.

## LIST OF FIGURES AND ABBREVIATIONS

### Figures

Figure 1	An illustration of RCT design.	p. 27
Figure 2	A schematic representation of psychological functioning.	p. 38
Figure 3	A simplified depiction of the process and brain structures involved in the perception and response to a potentially threatening situation.	p. 40
Figure 4	An illustration of potential treatment areas to be targeted for improving psychological functioning.	p. 44
Figure 5	An illustration of a well-balanced healthcare research design.	p. 48
Figure 6	Testing the effect of different contextual components.	p. 48

### Abbreviations

DBV	Dodo-bird verdict. Concept introduced in chapter 1 (p.5) and more thoroughly explained in paragraph 3.1 (p.14).
DSM	Diagnostics and Statistical Manual. See the American Psychiatric Association's (2014) website for more information.
ICD	International Classification of Diseases. See the World Health Organisation's (2015) website for more information.
ICF	International Classification of Functioning, Disability and Health. See the World Health Organisation's (2015) website for more information.
MPT	Music psychotherapy. See chapter 1 (p.6) for a definition.
MT	Music therapy. See chapter 1 (p.6) for a definition.
PT	Psychotherapy. See chapter 1 (p.6) for a definition.
RCT	Randomised Controlled Trial. See for instance Thyer (2001 <sup>b</sup> ) for an explanation.