Explaining the Underlying Psychological Factors of Consumer Behaviour with Artificial Neural Networks

By

Max N Greene

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

Marketing and Strategy Section of Cardiff Business School, Cardiff University

January 2016

Acknowledgements

I would like to express my sincere appreciation to my supervisors Professor Gorgon Foxall and Dr Peter Morgan for their continuous support and mentorship throughout these past years. I am indebted to them for numerous hours of discussion that helped formulate the ideas expressed here. They have always offered a tremendous encouragement for my research interest and helped my development as a research scientist through their invaluable advice.

I would like to thank the academic, administrative, and technical staff members of the Cardiff Business School who have always been there to help and advise in their respective roles.

I am very grateful to my friends and family for their continuous support and encouragement.

This research has been supported by the Economic and Social Research Council and by the Marketing and Strategy section at Cardiff University Business School.

Data supplied by TNS UK Limited. The use of TNS UK Ltd data in this work does not imply the endorsement of TNS UK Ltd in relation to the interpretation or analysis of the data. All errors and omissions remain the responsibility of the authors.

Summary

This thesis intends to advance our understanding of consumer behaviour, and proposes an extension to the theoretical and methodological framework of the Behavioural Perspective Model. Drawing on the intellectual tradition of connectionism and employing advanced artificial neural network modelling techniques, the research programme described here assesses the appropriateness of connectionist architectures in explaining consumer behaviour. This thesis traces the developments in the fields of consumer behaviour analysis to critically evaluate the significance of limitations inherited from radical behaviourism, and proposes a hybrid connectionist approach to address these methodological constraints.

The study is both highly quantitative and interpretative in nature, and generates a large body of empirical evidence to support the methodological and theoretical deliberations. Two types of data are used here: a simulated dataset to assess the capacity of the pruning algorithms to reveal the underlying relations within the data; and a consumer panel dataset to which the neural network algorithms are applied to develop predictive, descriptive, and interpretative connectionist models that aim to explain the consumer purchasing decision-making process.

Even though it is beyond the scope of this research project to propose mechanisms to explain all elements of consumer purchasing decision and it will therefore remain to be addressed as part of an ongoing collaborative research programme, the main conclusion to be drawn from this work is that the connectionist framework and artificial neural networks can be considered a significant contribution to the extension of theoretical and philosophical framework of intentional behaviourism.

Contents

List of Figures

1. Introduction

The study of consumer behaviour is an intricate and complex undertaking, and may often involve countless factors and variables that have an impact on the consumer and their physical and social environment: individuals, groups, and organisation participate in a number of processes with a purpose to select products, services, experiences, or even ideas that would satisfy their certain specific needs. Due to the nature of this complexity, it is a multidisciplinary endeavour that draws together elements of psychology, marketing, economics, and artificial sciences. The overall goal is of course to explain the consumer decision-making process, and provide a plausible model on an individual underlying level from both psychological and physiological point of view. Classical approach normally would involve a comprehensive interrogation of variables that may offer a degree of predictive and descriptive capacity to identify the level of linear relationship and significance they may exert over the ultimate consumer choice. Approaches that are more recent go a step further to employ advanced concepts of distributed representation to examine the consumer behaviour as an emergent process as a result of learning continuity. Nevertheless, consumer behaviour remains extremely difficult to predict and explain – this serves as a motivating factor to continue advancing the research in this direction and as a result continuously consider and critically review innovative methods and applications to extend the current body of knowledge.

The complexity of a subject matter that would benefit from an individual level of comprehensive analysis not only on the level of behaviour, but also on the level of the consumer learning and the mind and even the neurophysiological level information processing, often triggers the process of scientific decomposition of complex phenomena to study the comprising elements of the process independently – as a result, the scientific account is either fragmented and incomplete, or provides a variant of a *Black Box Model* where certain elements of the overall process are either assumed or otherwise effectively disregarded. To overcome this, a robust comprehensive theoretical and empirical framework to describe and explain consumer behaviour and the underlying psychological and physiological factors would be indispensable, and any progress in such direction would be not only essential to facilitate progress in the field of consumer behaviour, but could also be extended to wider context and provide a contribution to explanation and understanding such fundamental phenomena as learning, intelligence, and cognition – both human and artificial.

1.1 The thesis structure and contents

Chapter 1 introduces the research project and briefly outlines the structure and contents. The chapter provides a discussion around the overall motivation for the research project, and the focus of the research is succinctly discussed and summarised.

Chapters 2 and *3* demonstrate a wider context for the research project, which by the nature of the research questions addressed here touches upon a number of disciplines and research areas in the course of the inquiry as described in the chapters below, resulting in a multidisciplinary work that embraces the elements of critical behaviourism and cognitive sciences, traditional and neural networks modelling approaches, and theoretical frameworks that propose to extend the established theory around Behavioural Perspective Model with connectionist architectures. *Chapter 2* provides an overview of the field of consumer behaviour, discusses the theoretical and philosophical frameworks of radical behaviourism, and offers Behavioural Perspective Model. *Chapter 3* extends the discussion into the science of artificial, and introduces the field of artificial intelligence. The capacity of symbolic modelling methods are discussed at length and compared with the connectionist networks, offering a detailed account of neural network modelling techniques and architecture optimisation algorithms.

In *Chapter 4*, the research methods are explained in detail. The chapter provides an overview of the research questions and research methods employed as part of this research project. Here the research questions are proposed, followed by the discussion that aims to establish the philosophical position adopted here. Next, research method is outlined, where the data structures and variables are described. The modelling approach is explained and justified, and research process is outlined in a sequential manner.

Chapter 5 covers the analyses part of the project, providing a comprehensive account of the research methods employed. The statistical analyses employed throughout this research project are discussed in detail, and the specifics of the models developed in the course of research project are described. The testing procedure to support the line of inquiry is explained in detail, offering an overview of the results.

In *Chapter 6,* the results are discussed and interpreted within the wider context of consumer behaviour in general. The opening part of the discussion that revolved around the variable contribution and advanced connectionist modelling takes place here. The discussion is then extended to argue the appropriateness of connectionist modelling to provide the explanatory and interpretative account of consumer behaviour employing the pruning algorithms to optimise the network architecture to expose the core architecture. Theoretical implications are then discussed before the arguments are reviewed in the next chapter.

Chapter 7 offers a critical assessment of the research project and demonstrates precision, thoroughness, contribution, and comparison with its closest rival, the tradition of cognitive science.

In *chapter 8*, a number of possible future research directions are identified and briefly discussed: a number of possible strings of inquiry are identified, ranging from purely commercial applications to apply and test the methods proposed here in the industry, to highly theoretical and philosophical endeavours that would aim to explore the concept of distributed representation further and extend the line of inquiry into the field of swarm intelligence.

Finally, *Chapter 9* provides closing remarks, touching upon the contributions this research project aims to offer, and concludes with a summary of the research project.

2. Consumer behaviour

The interdisciplinary nature of academic marketing implies the tendency to adopt the perspectives and methodological techniques from the established fields such as economics and psychology rather than relying on the deliberately developed theoretical foundations. Thus, the philosophical and methodological assumptions tend to reflect the original body of inquiry and provide for marketing a misrepresentative and transitory theoretical foundation. In contrast, consumer behaviour analysis is concerned with phenomena central to marketing – the explanation of consumer choice – and offers a cohesive philosophical and theoretical foundation for the inquiry. Borrowing largely from the already widely adopted in academic marketing paradigms of cognitive psychology and behaviourism, the behaviour analysis provides a framework for explanation of consumer behaviour in its natural environment.

The research programme generates a body of knowledge concerned with the adequacy of radical behaviourism in explaining consumer choice, and involves advances in theory and philosophy of behaviour analysis and modelling of consumer behaviour, and offers means for consumer behaviour interpretation based on empirical research (Foxall, 2005). In doing so, research programme aims to determine the degree to which consumer behaviour can be sufficiently explained with radical behaviourism, and subsequently offers extension of theory from other fields of inquiry. Resulting efforts manifest in development of Behavioural Perspective Model (BPM) of consumer choice and lead to empirical research in consumer behaviour analysis and patterns of consumer choice; and offer novel ways for interpretation of consumer behaviour and extending the theoretical and philosophical base. Thus, it is possible to predict consumer behaviour to certain extent, and to demonstrate an insight into what contributing factors control it, but not so much to explain the behaviour beyond the identification of controlling stimulus conditions.

The conceptual framework developed in attempt to provide explanation of complex human behaviour is discussed in this chapter. Consumer behaviour in particular is the focus of this research project, which will draw upon the conceptual fields of behavioural economics, psychology, biology, and philosophy, aimed at building a unified interdisciplinary model of consumer choice.

As it may seem that a strictly behaviourist approach would not be able to provide a sufficient account of consumer behaviour on the individual consumer level, the concepts of *intentional behaviourism* and *super-personal cognitive psychology* are introduced (Foxall, 2004). This is further developed and explored in Understanding Consumer Choice (Foxall, 2005) and Interpreting Consumer Choice (Foxall, 2009), where consumption patterns are explored empirically employing the above mentioned concepts to expose the role of contextual-intentional psychology in everyday consumer decision making. In this chapter, the underlying philosophical assumptions are discussed in the context of theoretical and empirical aspects of consumer behaviour analysis.

2.1 Explanation of consumer behaviour

Even though some consumer behaviours can be identified in a rather obvious manner and marketing strategies are used by retailers in attempts to condition the customer, psychological approach of radical behaviourism has a lot to offer in terms of marketing concepts. It is necessary to explain the consumer behaviour employing the social and behavioural sciences and identify the deterministic factors that influence the decision environment. Radical behaviourism is primarily concerned with explanation of behaviour, assuming that behaviour is explained through environmental stimuli that predict the behaviour – the basic notion that carries not only the explanatory power, but also the comparative means to critically review and consider other methods of explaining consumer behaviour as a social, economic, and biological phenomenon. Thus, radical behaviourism can be viewed not as a sufficient element capable of providing a comprehensive account of human behaviour, but rather as an essential constituent in the theoretical and empirical pursuit of developing such system.

The research programme led by Foxall provides a critical review of radical behaviourism, while developing the theoretical framework that suggests the possibility of prediction, control, and explanation of behaviour in a process of contextualization of behaviourism in a broader scientific explanation. As a result of the research program, new theories of human behaviour have been introduced based on the principle of identifying the patterns of operant behaviour in the selective environment that shapes and maintains the behaviour while surpassing the theoretical constraints of radical behaviourism: *intentional behaviourism* and *super-personal cognitive psychology*.

The focus of the research programme in early years revolved around the theoretical developments consisting of critique of predominant at the time cognitive view of consumer research from the radical behaviourist perspective, followed by establishing the basis for the alternative interpretation of consumer behaviour – *Behavioural Perspective Model* (Foxall, 1990). Subsequently, the empirical investigation was undertaken to substantiate the proposed argument that establishes behaviourist interpretation of consumer choice as a powerful alternative to cognitive theory and other non-behaviourist reasoning; and provides an understanding what behaviourism and other approaches are able to offer exclusively, identifying not only the inadequacies and limitations of radical behaviourism that need be supplemented, but also creating a basis for evaluating the alternative non-behaviourist approaches.

Research programme aims to establish the means by which reliable scientific interpretation of consumer choice is plausible – but one of multiple interpretations in the relativist sense, all subsequently tested with scientific method with the purpose of producing a comprehensive body of knowledge.

2.2 Intentional Behaviourism explained

It is important to outline the methodology of Intentional Behaviourism research programme before its elements are explained in detail in the following sections, which comprises three conceptual stages as shown in Figure 1 (Foxall, 2016).

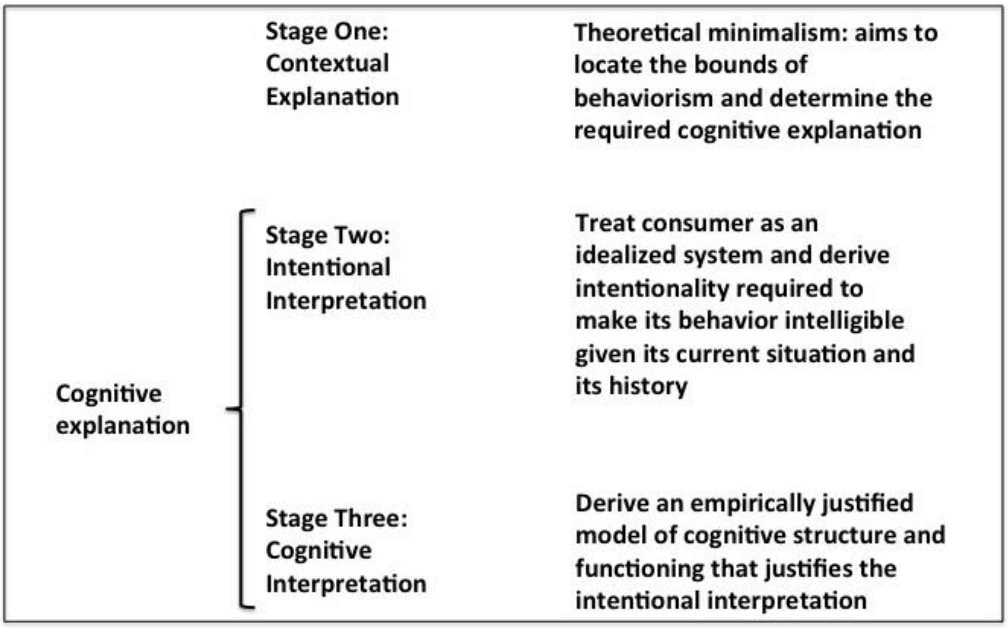

Figure 1. Methodology of Intentional Behaviourism.

Stage One is *Theoretical Minimalism/Contextual Explanation*, and is primarily associated with the extensional behavioural science used to delineate the scope of behaviourist explanation. Choice is represented at this stage as selection amongst alternatives by carrying out one type of behaviour as a proportion of all behaviour instances. Reliant on extensional explanation and three-term contingency of radical behaviourism, behaviour is evaluated empirically employing experimental and statistical methods by deconstructing observed behaviour and identifying factors responsible for prediction and control.

The purpose here is twofold: as part of delineation of behaviourist explanation of consumer choice, it is identified which particular aspects of behaviour are inadequately explained by behaviourism, and determined what form would be required for the intentional explanation to interpret the behaviour. Often stimulus conditions required for prediction and control of behaviour are inaccessible, and a number of behaviourist limitations become apparent: inability to account for continuity of behaviour, absence of personal lever of explanation, and delineation of behaviourist explanation. Continuity of behaviour issue in behaviourism occurs when antecedent and consequent stimuli normally used to explain behaviour with *n*-term contingency are inaccessible with interpersonal level of explanation. Personal level of explanation is a behaviourist response to phenomenological subjectivism aimed to provide a behaviourist account for thoughts, feelings, and other private intentional constructs, which cannot be explained with extensional terms. Delineation of behaviourism would aim to establish the scope and limits for the behaviourist approach to explain behaviour (Foxall, 2016).

Stage Two is the *Intentional Interpretation*, where consumer is treated as an intentional system and attributed with the intentionality to maximise utility relying on the learning history within a given behaviour setting. While account of behaviour here attributes a set of thoughts and emotions to the consumer as part of the intentional interpretation, this is consistent with the results observed as part of the empirical modelling programme in Theoretical Minimalism stage. The possibility to refer to this empirically grounded foundation serves as a constraining element for the intentional interpretation of consumer behaviour to restrain psychological speculation (Foxall, 2016).

In Stage Three, *Cognitive Interpretation*, empirically supported model of the underlying cognitive structure and functioning is proposed, consistent with the intentional interpretation account of behaviour offered in Stage Two. The aim here is twofold: develop a micro-cognitive psychological construct consistent with both intentional interpretation and sub-personal level of consideration in the neuroscientific sense; and define a macro-cognitive psychology that demonstrates the consistency of intentional interpretation with super-person level as studied by behavioural scientists (Foxall, 2016).

Although intentionality is not ascribed to sub-personal and super-personal levels of explanation, it is an essential point of Interpretative Behaviourism that intentional interpretation must be corroborated and supported by the extensional account of behaviour typified by empirical scientific method of radical behaviourism and neuroscience (Foxall, 2016).

The elements of Intentional Behaviourism will be discussed in detail in the following sections.

2.3 Radical behaviourism

Radical behaviourism, the metatheory of behaviour analysis, is established on the principle that objective and empirical methods of natural sciences can be applied to the analysis of human behaviour; and states that the behaviour is explained when the environmental factors that influence the rate of repeat behaviour are identified, and response can be predicted and controlled through the manipulation of reinforcement contingencies. Notice that no causal reference is made to the internal states or processes or events such as mood or intention or personality traits as is common in cognitive theories – they are not ignored in behaviourism but rather are classed as responses that require a separate explanation, rejecting the incomplete theoretic development reliant on mental or conceptual entities that can be said to reside at other than observed behavioural level. Such theories can be considered *incomplete* as they fail to identify the factors that account for internal processes and events such as environmental precursors causal to behaviour; *fictional* as they tend to infer the internal causes from the behaviours they are supposed to explain; and *superfluous* as behaviourism provides a simpler explanation of behaviour through environmental factors that control behaviour without relying on fictional inference.

In behaviour analysis, the stimuli that are said to control behaviour need be explicitly described and related to the rate of response. In the laboratory setting of experimental operant psychology with animal subjects, it is possible to establish explicitly the discriminative and reinforcing stimuli and their causal relationships with response rate in pursuit to identify the elements of controlling contingencies, and predict response rates according to the reinforcement schedules, thus explaining simple behaviours by reference to their environmental antecedents and consequences. The three-term contingency is able to succinctly describe the causal mechanisms of behaviour analysis: the (1) *discriminative* *stimuli* in the presence of which (2) the *response* is emitted, and the (3) *reinforcing* or *punishing consequence* produced as a result.

In the context of complex human social interaction however, unlike the laboratory setting, it may often be impossible to isolate environmental contingencies controlling behaviour with any degree of certainty, resorting as a result to conceptual extrapolation of learning process derived from animal operant conditioning studies. Experimental laboratory operant analysis is said to provide the scientific *explanation* of behaviour, which is then extended to suggest not an explanation but rather a plausible *interpretation* of more complex behaviour situations.

2.3.1 Interpretative Behaviour Analysis

Skinner (reprinted in 2014) argues that even though it may not be possible to determine the contingencies that control response rate in complex behaviours with any degree of accuracy and precision comparable to laboratory experimentation, it is feasible to offer a plausible account of complex behaviours, such as verbal behaviour, based on the extension of the scientific laws formulated during the analysis of smaller decomposed problems. The behaviour analytic interpretation, even if unverifiable experimentally, is preferable nonetheless to the interpretations that are not supported by the experimental work on smaller decomposed phenomena. Astronomy could be used to illustrate the same principle, where main source of information about celestial bodies and other objects is electromagnetic radiation and numerical models are employed to reveal the existence of phenomena and effects otherwise unobservable. Thus, the extension of learning principles of human and non-human operant responses derived in the scientific laboratory setting is employed in the extended behaviour analytic account (derived from the scientific knowledge) of human behaviour in complex social situations.

2.3.2 Plausibility

During the process of interpretation, operant analysis is inevitably altered – behaviourism speaks of *plausibility* in terms of persuasive power of the interpretative account thus offered referring to the larger body of experimental inquiry to support the claim, where interpretation is not equal to extrapolation. The consumer behaviour model is then assessed on the explanatory plausibility of the variables to provide a persuasive account that demonstrates behaviour patterns, and the empirical correspondence and ability to derive operational variables useful in further investigation. The plausibility of interpretative account relies on its empirical correspondence with the objectively acquired information. All science relies on interpretations when explanation is no possible, as it often dictates the route of investigation and explanation, and is an inherent component in synthesis and amalgamation of information.

As compared with operant behaviourism that predominantly specializes in simple observable behaviours with empirically identifiable determinants such as pigeon pecking, interpretative accounts of complexity based on the findings from simpler scientific experiments offer qualitatively different type of knowledge. This fundamental difference explains why basic research objectives and technology requirements set forth by behaviourists – prediction and control in particular – are not the sole focal point while extending the operant principles of explanation in the form of interpretative plausible account.

The BPM is able to demonstrate that only minimal modifications of the basic behaviour analytical model are required to offer a plausible interpretation of complex behaviours within a critically derived behaviour analytic framework and only deviating from radical behaviourism in a manner of extending it based on logical criticism and empirical evidence. One of the basic guiding principles of BPM is the proposition that human behaviour in any setting can be operant, where the continuity of environmental influence in a variety of situations is a fundamental assumption.

2.4 The Behavioural Perspective Model

The Behavioural Perspective Model extends the framework of behaviour analysis by recognising the discriminative stimuli and reinforcers as independent variables that determine the response schedule, and relating them to the rate of response in a purchasing decision setting as a dependent variable – as a result, the model is able to provide an interpretation of complex social consumer behaviour situation in a complementary to other alternative approaches (predominantly cognitive) manner by considering the ontological and methodological aspects. BPM is ultimately aimed to develop a comprehensive account of consumer behaviour that would incorporate a synthesis of both cognitive and behavioural explanations in two ways: (1) by providing a plausible reference to conceptualization of situational influences on consumer behaviour, and (2) suggesting novel consideration of marketing strategy. Cognitive decision science assumes a purchasing decision to be a result of goal-directed process where consumers deliberately plan the course of action and utilize resources to acquire a desired benefit, the process during which external influences on consumer choice are not sufficiently taken into consideration and decision setting inadvertently decontextualized as a result. Those scarce attempts to explain consumer behaviour in terms of external stimuli suggest the prospect of developing a unified framework, but fail to propose a model of purchasing decision-making that is both based on empirical foundation of operant psychology and is relevant to marketing and strategy. Theoretical and conceptual framework of BPM is developed to address these concerns. Moreover, the model associates the previously identified and described patterns of behaviour with appropriate contingencies in a reliable manner, offering behavioural interpretation of consumer behaviour that is not postulated in conflict with alternative theories of explanation, and embracing the multiplicity of explanatory mechanisms in a relativist sense. Current research programme is tasked with the synthesis of intrapersonal and environmental causes of behaviour, where the relationship between the utilitarian and informational reinforcement and the affective cognitive theory is contemplated. The applied nature of BPM offers considerable benefit to the field of marketing research as the inherent categorization of contingencies that influence consumer decision-making and explanatory variables are practically applied in customer-centric marketing strategy.

Operant behaviourism as a school of psychology, as opposed to cognitivism and phenomenology, is better defined in terms of coherent philosophy of science, subject matter, methodology, and explanatory power – the notion that suggests operant behaviourism as a particularly well suited theoretical approach for the field of consumer research. Moreover, it has been shown that economic behaviour of animals is operant; and some consumer researchers have considered the possibility of employing the behaviour analysis to purchasing decision-making and the process of consumption. It is unclear therefore why operant behaviourism is not commonly used in explaining consumption in terms of situational and environmental influences, and why the comprehensive theoretical perspective of consumer psychology able to deal with the situational effects on choice has never evolved despite the research on consumer task orientation, temporal perspective, antecedent states, and physical and social consumer surroundings. Several factors that contribute to attribution of oversimplified nature to behaviour analysis may be responsible in this matter.

Firstly, given the interdisciplinary relativist nature of consumer psychology, it may seem odd to have behaviourism excluded as one possible explanation of behaviour because of the misattributed idea that behaviourism was superseded by the cognitive paradigm due to inherent ontological restrictions of behaviourism proving it inadequate as a mental discipline. On the contrary, behaviourism is able to provide a philosophical outlook from which other approaches may be critiqued, and encourage empirical data otherwise unavailable to be generated – as recognized by consumer behaviour researchers. Secondly, researchers have largely failed to account for the great difference between human and non-human cognitive capacity and the human ability to create and adopt rules that describe contingencies of reinforcement while extrapolating from the general findings with animals to human consumer decision-making, assuming unwarranted proportion of continuity between animal and human behaviour. Thirdly, the overall complexity of the human consumption environment and the non-price elements of marketing have been largely overlooked, focusing instead almost exclusively on the effects of price on demand. Quite the opposite, in advanced modern economies demand is contrived and deliberately created by most of the marketing effort in a consumption environment that contains the vast amount of choice alternatives available to human consumers. Finally, it is a common practise in marketing and related disciplines to disregard the philosophical and explanatory implications of behaviourism, rather directing research efforts towards the use of reinforcement schedules to increase the rate of retail purchasing, often forgetting that behaviourism is not the science of human behaviour, but the philosophy of the science of human behaviour. Thus, BPM is set to address these issues.

Unlike marketing and consumer psychology that merely draw upon certain aspects of behaviour analysis, the objective of BPM research programme is to develop a critical understanding of consumption by developing as complete a model of consumer behaviour based on behaviour analytical method as is reasonably possible, establishing in the process the extent to which behaviour analysis alone is able to explain consumer behaviour, and determining where it may be practical to modify and extend the framework of analysis. Research programme is also concerned with the evaluation of the ability of behaviourist account to explain complex consumer behaviours, and its contribution to the advancement of consumer psychology. In doing so, BPM distinguishes relatively closed consumer behaviour setting where the environment is similar to operant conditioning, and relatively open consumer behaviour setting where the environment is rich with alternative to operant conditioning explanations of behaviour. Furthermore, the model recognizes not only utilitarian type of reinforcement such as pleasurable and utilitarian consequence of behaviour, but also informational reinforcement that can take a form of feedback from other members of social system for example. And most importantly, the behaviours are attributed to proximal latent internal causal elements such as verbal discriminative stimuli in covert rules of behaviour and reinforcement contingencies; as well as environmental determinants of consumer behaviour that require both analysis and interpretation. Thus, the BPM aims to assess the adequacy of behaviour analysis to provide a rigorous scientific account of purchasing and consumption phenomena in a complex setting with multiple sources of causation.

2.4.1 The framework of Behavioural Perspective Model

Patterns of consumer choice are related to the differing environmental consequences in BPM (Foxall, 1990, 2004, 2005, 2009), proposing three kinds of consumer behaviour consequence: (1) utilitarian reinforcement representative of benefits from buying or owning the product, (2) informational reinforcement represented by the social aspects of consumption, and (3) aversive consequence posited by such events as relinquishing money and product alternatives (see Figure 2). Antecedent events comprised of any physical, social, or temporal elements that serve as signals for potential consequence form the behaviour setting continuum capable of either facilitating or inhibiting the consumer choice, ranging from open setting that offers great freedom for a consumer to act and choose to a closed setting where consumer behaviour is largely dictated by other than consumer agents. The consumer is represented through their learning history that takes into account the aggregate consequence of previous behaviours and the present state of the consumer.

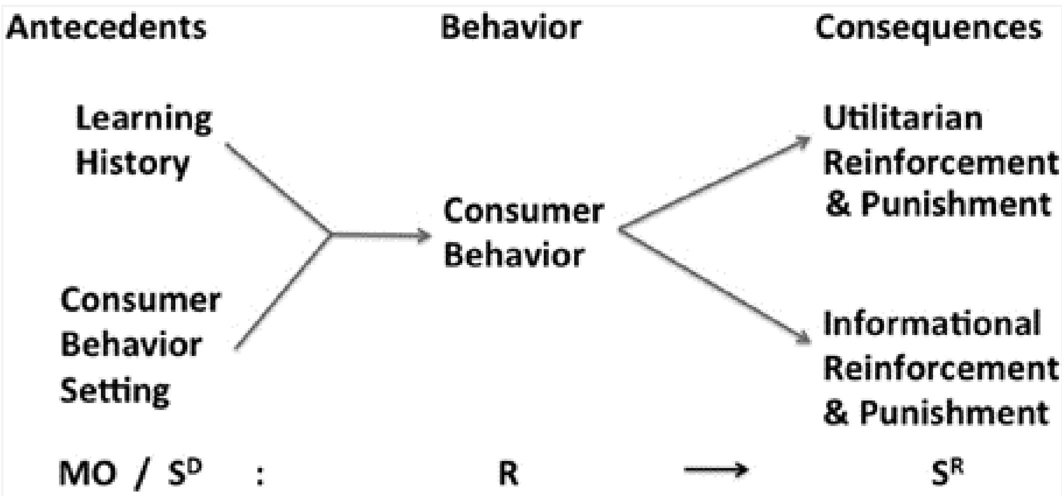

Figure 2. The Behavioural Perspective Model.

The combination of high/low utilitarian (UR) and informational reinforcement (IR) identifies four distinct classes of consumer behaviour (see Figure 3): (1) low utilitarian and informational reinforcement constitutes *Maintenance* and involves the satisfaction of basic needs, (2) low utilitarian reinforcement and high informational reinforcement constitutes *Accumulation* such as saving or collecting, (3) high utilitarian reinforcement low informational reinforcement constitutes *Hedonism* such as consumption of popular product, and (4) high utilitarian and informational reinforcement represents *Accomplishment* reflective of social and economic achievement (Foxall, 2009). The addition of behaviour setting continuum provides eight contingency categories as a functional consumer behaviour analysis and as means of interpreting factors of consumer behaviour.

	High Utilitarian reinforcement	Low Utilitarian reinforcement
High Informational reinforcement	Accomplishment	Accumulation
Low Reinforcement reinforcement	Hedonism	Maintenance

Figure 3. Operant Classes of Consumer Behaviour.

2.4.2 Brand choice

Contrary to traditional marketing literature that assumes consumers to explore the entire spectrum of choice, majority of consumers tend to exhibit multi-brand purchasing patterns from a small repertoire of preferred brands, a subset of the entire range. Even though previous research describes consumer behaviour in terms of determinants of patterns, few offer discussions on the consumer goals or the underlying factors that influence the consumer decision-making. Experimental behaviour analysis however is not only able to demonstrate that consumer choice adheres to the patterns proposed by behaviour analysis and behavioural economics, but also offer a possible theoretical extension of operant psychology.

2.4.3 Verbal and affective response

Behaviour analysis has traditionally relied primarily on the idea of schedules of reinforcement – something that may impose extreme difficulty of implementations in an open market environment. Alternatively, interpreting complex consumer behaviour in terms of reinforcement pattern can be demonstrated with verbal and emotional responses to a consumer situation. Each of the eight contingency categories was shown to correspond with a certain combination of basic emotional response: pleasure with utilitarian reinforcement, arousal with informational reinforcement, and dominance with behaviour setting continuum.

2.4.4 Attitudinal-behavioural consistency

Psychological research on attitudes may actually be interpreted as behavioural, where observed variables often are verbal behaviour in nature: measures of attitude and intent tend to be situation-specific, emphasizing contextual determinants of behaviour as a result rather than showing the cognitive causes of behaviour. Previous history of behaviour holds a significant role – being the best predictor of current or future behaviour, and the adherence to the social cognitive paradigm – the two aspects that pose a difficulty for the attitudinal researchers to examine the continuity between the environmental factors and the verbal and non-verbal behaviour response. For these reasons, recent findings in attitudinal psychology may support behavioural rather than cognitive model for human behaviour.

Behaviour analytic approach states that initial behaviour is rule-based, and subsequently comes under the contingency control. Foxall argues that the process follows a somewhat more complex path, suggesting that initially consumer may have no previous experience with a new product or brand. Rules associated with other similar product or situation may be utilized however, which are deliberately assessed in terms of compatibility and accumulated experience before being adopted by the consumer. Repeat successive behaviour and accumulation of experience will generate rules that will guide behaviour, replacing the deliberation process, and transforming it from rule-governed to contingency-based. The difficulty lies in illuminating this transitional process without reliance on the implied theoretical structures.

2.5 Intentional psychology

As part of any discussion on behaviourism as a psychological approach, it would be fundamental to review how the language is used by its practitioners, and why the discussion on intensionality may not be simply circumvented while the usage of intensional language is interrelated with intentional explanation.

2.5.1 Intentionality

Social cognitive psychology argues that consumer decision-making is guided by *preferences*, *likes*, *wants,* and *needs*, and has a *positive attitude* or *intent* towards the purchase – the level of explanation that relies on employing the mentalistic terms that would not suffice if the aim is to go beyond the cognition-behaviour approach and contemplate the philosophical basis for the explanation of consumer behaviour. The mentalistic terms used above are all *intensionalistic* in a way that they offer a qualitatively different type of explanation, as they tend to implicitly refer to something outside themselves: *prefer something* rather than just simply *prefer*, *want something* rather than just *want*, and so on. Quite the reverse, the words that are not essentially mentalistic such as *run* do not rely on the preposition that follows them to provide the precise meaning (i.e. prefer *this product* to *that one*): it is never *run something* but rather just *run*. Precisely this *intensional* use of language where what follows the proposition (i.e. *intensional operators* that added to extensional statement to produce intensional statements) cannot be substituted with an equal in meaning term without changing the truth-value of the statement is what behaviourists would prefer to avoid in explaining behaviour. This is paramount in extending behavioural science, as *intensional* language is not reducible to *extensional* language through paraphrasing or other means, thus making the extensional account not possible. When describing inherently intentional phenomena though such as *perceiving* or *believing*, it is inevitable that intensional language is being used, and therefore intentional explanation is being used – suggesting that in addition to extensional behavioural science an intentional explanation may be unavoidable to the exposition of science.

It is possible to argue that radical behaviourism may not be able to fulfil all the essential requirements central to explanation of behaviour, and behaviour deals on a personal level of inquiry, provides adequate account of behaviour continuity, and demonstrates that explanations of complex human behaviour can be sufficiently attributed to reinforcement stimuli.

Radical behaviourism strives to provide account of behaviour in a strictly extensional manner, distinguishing itself from cognitivism by deliberate aversion of intensional language. The scientific method of behaviourism, the behaviour analysis, is focused on prediction and control of behaviour in terms of environmental consequences and antecedent stimuli. Assuming behaviour is environmentally determined in terms of learning history and reinforcing or punishing consequences, behaviourism is able to provide an undeniable capacity for explanation of behaviour – it does not necessarily dictate however that further conceptualization need be confined to the behaviourist philosophy.

Personal level refers to human activities and sensations rather than sub-personal physiological processes in the brain and nervous system, and comprises of first-person (intra-personal subjective experiences) and third-person (inter-personal objective experiences) perspectives. Extensional radical behaviourism does not consider either one in any adequate manner: first-person perspective necessitates the use of intensional language to provide a comprehensive account of certain behaviours, and third-person perspective is not considered qualitatively distinct and is explained in terms of first-person perspective.

Continuity of behaviour is central to radical behaviourist theory as it is closely interlinked with the concept of behaviour reinforcement and continuous learning history formulation. What exactly is learned during the learning process however is not evident from the behaviourist theory, as it may be difficult to identify the required elements necessary for the learning to occur (discriminative stimulus, response and reinforcing stimulus), at which point it is presumed that something does happen within the individual (likely on physiological level) that satisfies the behaviour continuity requirement.

Thus, the limitations of radical behaviourism to provide a plausible interpretative account of behaviour in a variety of situations in a manner satisfactory of the rigorous scientific standard require intentionalistic supplementation to extend its significance.

If radical behaviourist theory were to be extended to consider intentionality, it would require appropriate definition in extensional terms. One way to do this is to distinguish two qualitatively separate levels: the physiological level that applies to neurological networks and processing, and the level of intentionalistic interpretation to which intentional abstractions like beliefs are attributed. Whereas the physiological level is commonly well understood, it cannot be said the same about the intentionalistic level, as the conceptual description is circular at best (i.e. pain is painful). Thus, the intentional level should not be seen as providing additional characteristic of phenomena, but rather a qualitatively different to the extensional explanation manner of interpretation of the same process or event.

2.5.2 Philosophy of intentionality

Philosophy refers to intentionality as capacity of the mind where mental states refer to or are about something other than themselves.

Two distinct types of intentionality may be recognized: the *intrinsic intentionality* that exists in a human mind possesses the ability to transfer its inherent intentionality into an object (so called representational artefact), at which point the *derived intentionality* emerges. Even though the intentionality may seem to be stripped of its analytic ability in the process as seemingly all objects are about something, the derived intentionality is nevertheless dependent on the original intrinsic intentionality. Thus, the private intrinsic intentionality relates to the concept of the subjective mind and includes the experiences of believing and desiring, and therefore making it possible to attribute these experiences to a third person only because of experiencing them ourselves.

The subjective experiential level is dependent on the inherent properties of the conscious phenomena: the consciousness that includes the emotions and thinking or cognition, the ontological subjectivity that specifies the existence of a conscious state only if experienced by the subject, and unity which is required for various aspects of consciousness to function in a collective manner to produce the experience of the situation.

Conscious experience could be comparable with Skinnerian private (covert) behaviour, but may also be internally augmented by interpretation in addition to being externally determined by contingencies. Consequently, different language from extensional behaviourism should be used as private (covert) behaviour contains intentional behaviour such as believing.

Speaking of intentionality in terms of objective approach (third-person), it is necessary to consider what object may be considered as being or having a mind – something that theory of mind considers. Intentional behaviourism necessitates only the first-order intentionality. The need for intentional stance all together must be considered as well, as the concept of mind is relevant only while speaking about human persons where physical, design, or contextual stance may prove insufficient; where the experience includes not only the physiological and behavioural aspects, and the mind is more than just the brain but also includes the subjective consciousness and the cognitive awareness. This creates the setting for the first-person experiences such as thoughts and beliefs, and the third-person experiences are inferable from analysis of the physiological and behavioural factors associated with such experiences.

2.5.3 Intentional psychology

Content does not occur within the neural event itself, but rather is a supplementary interpretation of such event, a justification for registering certain local content on the personal level. Intentionality describes what extensional theory is able to describe, but in a different manner. Thus, the sub-personal extensional physiology and the personal intentional explanations belong to qualitatively different distinct content levels. Content is added by the evolutionary logic principle: findings produced from the natural selection process on the physiological level must allow for explanation and prediction of behaviour on the intentional (personal) level.

Dennett (1981) makes a distinction between the three kinds of intentional psychology: (1) *folk psychology*, (2) *intentional systems theory*, and (3) *sub-personal cognitive psychology*. *Folk psychology* revolves around the causal theory of behaviour and presumably provides the source for the other two. *Intentional systems theory* develops the belief and desire aspects of folk psychology to predict and explain behaviour semantically on the personal level, but largely ignores the internal structure of the complex intentional system. On the contrary, *sub-personal cognitive psychology* deals with the syntactic explanation of the brain function. The underlying internal structure is required to provide an explanation for components of the *informational systems theory* to predict systemic behaviour on the personal level – something that is the primary task of *sub-personal cognitive psychology* as it identifies the constraints of design and clarifies how personal systems excel in intentional systems. This will be further discussed in the sections below.

2.6 Intentional behaviourism

In this section, the two prominent accounts of behaviour – the *radical behaviourism* and the *intentional psychology* – are discussed in attempt to provide an overarching theoretical framework that would incorporate the two traditionally opposing fields in a unifying psychological model of *intentional behaviourism*.

The vocabulary of behaviour theory is constrained by design to include purely behaviouristic terms to describe and explain behaviour, and behaviourists are deemed to face a considerable difficulty while trying to accommodate the thinking of cognitive psychologists and therefore tend to adopt one of the two routes: either incorporate the language of intentionality, which would inevitably lead to means of explanation detached from behaviourist method; or more often prefer to stay with the restricted behaviourist vocabulary, and as a result limit the range of explanation available to behaviourists, effectively restraining the potential level of contribution to the wider discipline of psychology. The beliefs and desires assume a central role in intentional psychology and serve as a base for cognitive psychology, whereas the opposing behaviourist theory strictly excludes any intensional terms deemed unable to provide reasonable means of explaining behaviour. As a consequence, while critical behaviourism is able to provide predictive account of behaviour, it struggles to address within behaviourist terms such essential notions as *personal level* of analysis and *continuity of behaviour* – which generally tend to be either ignored altogether or inevitably explained with intensionalist terms. The following paragraphs will discuss this in detail.

2.6.1 Cognitive psychology and radical behaviourism

Before the discussion and critique of theoretical and ontological specificities of radical behaviourism can be continued, first of all it is important to acknowledge what can be effectively seen as a dualist nature of the current state of the field that makes the definitions and any subsequent discussions difficult: the radical behaviourism that can be associated with Skinner (1938, 2014) as a central figure in refining and expanding the paradigm, as opposed to those who undertake an active role in extending the discipline into the realm of intellectual inquiry to provide a more comprehensive account of behaviour (for example Rachlin, 1994). Even though those behaviourists that belong to the latter category remain with the extensional mode of explanation, it is often the case that they go beyond the precisely defined constraints of radical behaviourism as described by Skinner, and operate past the experimental laboratory setting and in the realm of interpretation and theory development. Without explicitly explaining the extent of deviation from the Skinnerian radical behaviourism, the explanation leads to wider implications as a result of inevitably adopting new forms of language that belongs to intentional systems of explanation. Thus, the two fields are taken to represent the incommensurably opposing views in explaining the phenomenon of behaviour, and overarching conceptual framework relying on both radical behaviourist and cognitivist modes of interpretation would be required to provide a more robust and all-embracing account of behaviour – *intentional behaviourism* (Foxall, 2009).

The essential difference in explaining behaviour employing cognitive and behavioural approach is linguistic in nature: whereas radical behaviourism unsympathetically evades the use of intensional idioms as explanatory means, intentional explanation is inherently imbedded in the underlying philosophical structure of cognitivist theory. The reason for avoidance of intensional idioms in radical behaviourism is that the extensional and intensional sentences are fundamentally different, and since radical behaviourism firmly relies on the use of extensional knowledge by definition, the adoption of intensional language to explain behaviour warrants the supposition that the very explanatory mode of the researcher has effectively gone astray from the radical behaviourist paradigm. Moreover, the difference between the extensional and intensional sentence types is more fundamental than that – it is not possible to simply translate one type of sentence into another type, as reducing intensional language to extensional would inevitably require adding additional information in the process to maintain the meaning and the truth value of the sentence unchanged. If intensional language is employed by researchers to explain behaviour, it is by definition a method of explanation that refers to some other than radical behaviourism theoretical framework. Extensional language, as opposed to intensional, does not contain intensional terms and is referentially transparent in a way that it allows the substitution of any expression within the sentence to be replaced by another with same extension without changing the truth-value of the sentence: for example, it may be true that Jones believes that the train to London has arrived, but not that Jones believes that the train number 128 has arrived, even though the train to London is in fact the train number 128 and the two have the same extension. Nevertheless, the problem lies in the fact that these two expressions have two different intensions and therefore different subjective meanings, and if the truth-value of the sentence about Jones to be preserved, the two are not substitutable. This contradicts with the normal use of language for scientific expression, as the use if intensionality presupposes a certain degree of subjectivity, and intensional idioms allow themselves to be understood only in each other's terms without the possibility to circumvent the use of intensional vocabulary by any other means save entirely abandoning the use of intensional language altogether (Quine, Churchland, & Follesdal, 2013).

2.6.1.1 All-inclusive explanation of behaviour

When considering complex behaviours however, it becomes readily apparent that while radical behaviourism is exceptionally good at controlling and successfully identifying the environmental events to predict behaviour, the explanation (in a more general sense of a word, not as it is understood and defined by radical behaviourists) of behaviour it is able to provide could be insufficient in a number of ways. Radical behaviourism is unable to provide an adequate account of (1) behaviour on a personal level in addition to behaviour-environment relation, (2) continuity of behaviour, and (3) delimitation of behaviourist interpretation. This of course does not require a cardinal change of radical behaviourism – on the contrary, the field must continue advancing the behaviourist programme within the paradigm and provide a robust predictive account of how behaviour can be determined by its consequences. Nonetheless, it should be continuously challenged by another theoretical approach to identify the apparent constraints which radical behaviourism imposes on the explanation of behaviour in attempt to develop a more robust all-inclusive explanation by incorporating useful constructs from cognitivist and intentional psychology. The three areas that draw attention to the limitations of radical behaviourism are briefly discussed in the following paragraphs.

The notion of *personal level* explanation of behaviour is twofold in radical behaviourism: the matter of first- and third-personal sentences analysis. Talking about first- and third-personal sentence structure, in the practice of radical behaviourism both types can be and should be analysed and interpreted in the same manner following the uniform process of inquiry – first-personal verbal behaviour for example is just behaviour to be explained using control variables at the time of occurrence, and the verbal behaviour is not seen as a reference to something but rather is explained within the history of context in which it happens to occur. However, when one says using intentional language, "I can't find my keys" – the statement cannot be translated into extensional language, as inevitably new information will have to be provided. Statements like "I am looking for my keys" or "I am not succeeding at the task of trying to find my keys" do not carry identical meaning to the statement "I can't find my keys" as the personal subjective dimension that adds to the explanation of the behaviour associated with the experience is unverifiable externally. “I can’t find my keys” in fact is the only unique description of the behaviour, whether it can be observed to include the tasks usually performed by the persons who are trying to find their keys so they can leave, or they are just trying to find their keys to put them in the correct place to make sure they can leave as soon as they need to in the future and can avoid going through this process in the inconvenient time, or perhaps even express a general desire that somebody else should perform the task of looking for their keys for them – something that only that particular person can be said to *just know* about their behaviour. This *first-personal knowing* is different to a private event as described in radical behaviourist doctrine: although equally contingent on learning history, first-personal knowing is not a result of analysis of the learning history, but rather an intentional statement that is a product of personal experience and therefore is outside the realm of scientific inquiry; and translating intentional sentence into extensional would not be possible without adding new descriptive value to the statement (or losing certain descriptive elements) in the process.

Continuity of behaviour limitation attributable to radical behaviourism could be discussed touching upon a number of points. First, even though it can be said that intentionality is unable to provide a conclusive account of continuity of behaviour either, it becomes apparent that in a less controlled setting interpretation rather than experimental laboratory-type work becomes more essential. It is at this point, where intentional language is being adopted by some radical behaviourists

– perhaps unknowingly – in attempt to broaden the scope and improve the explanatory account. Second, in practical application, it is extremely difficult to keep good track of the history of reinforcing contingencies even in a controlled laboratory setting – basically impossible in the case with human decision-makers, questioning the possibility to carry out a behavioural analysis to the full extent altogether without eventual contribution and borrowing from cognitive psychology. Third, it has been proposed that eventual physiological findings would be able to provide a mechanism that would reveal why the process of occurrence of behavioural continuity – the notion which goes against the asserted stand-alone non-reliant on other scientific disciplines nature of radical behaviourism. It is not to say however that this behaviour is independent in its entirety from the external contingencies – what this means is that there is no legitimate way to understand and explain this behaviour other than using intentional terms; this *knowing* is intentional, and so is the explanation and the expression of it – something that cannot be accounted for within the terms of radical behaviourist paradigm. There is an occasional remark in behaviourist literature however, that describes the tendency for respondents to develop subjective rules on a personal level during operant experimental work that render them insensitive to the varying levels of reinforcement. Naturally, such interpretations cannot proceed without incorporating intentional terms to describe private events such as thoughts, and would inevitably require going beyond the delimiting mode of explanation defined by radical behaviourism – as is the case with the discussion of self-rule formulation for example, where the description could only be expressed in intentional terms of the individual.

Finally, the topic of delimitation of radical behaviourist interpretation needs to be discussed, as it becomes increasingly difficult to employ the thee-term contingency to develop a comprehensive account of behaviour beyond the laboratory setting: the plausibility commonly taken to be sufficient for radical behaviourist account of interpretative research is not sufficient to meet the claims of validity and reliability when applied to complex human behaviour. Rachlin (1994) proposes the concept of teleological behaviourism, where complex behaviours are interpreted in terms of final consequences of behaviour: one might be looking for the keys to leave the house – and to continue, to go to work, to earn the salary, to provide for the family, to be a good father, and so on. It should be obvious that considering behaviour as such a sequence while disallowing intentionality would immediately pose an unsolvable problem in a number of ways. Primarily, the necessity to examine the whole sequence of consequences to provide a comprehensive explanation of behaviour may very well be practically impossible, as it will have to be extended indefinitely. Furthermore, even if the final event offering the ultimate cause could be identified and the complete sequence of consequences examined and analysed, it would nevertheless be an unconvincing method to provide an explanatory account for the original behaviour which may be so remote from the final cause that no empirical event of the original behaviour had any contact with the final ultimate cause. In addition, the very notion of deliberately developing the sequences to arrive at the final cause relies on such inherently intentionalistic concepts as rationality, optimisation, maximisation, and others – otherwise what else is there that would prevent the sequence to be developed instead as one of the following scenarios: looking for the keys to put them in the fridge, or to throw them into the rainwater drainage, or any other scenario that does not follow one of the rational in one way or another and intentional in nature motivation?

2.6.2 Intentionality and behaviourism

As discussed above, the incompleteness of radical behaviourism is dependent on the prescribed confinement to exclusively using the extensional language, which provides limited benefits beyond the laboratory settings to explain the observed behaviour: for example, it is very difficult to say anything more than the basics in terms of contingencies of reinforcement about behaviours which are not sensitive to schedules of operation without employing the concepts of private events. Avoidance of intensional language in radical behaviourism strengthens operant class, but as a result only able to provide a description rather than an explanation of generalizations, and as such could only be sufficient for prediction and control, but not for a comprehensive understanding of behaviour. In addition, it is necessary to consider other modes of explanation if comprehensive explanation of behaviour is not possible in extensional behaviourist terms – modes such as intentionality and intentional idioms. Nevertheless, complex human behaviour can be explained not only in terms of contingency shaping, but also in terms of rule-governance.

By incorporating the language of rule-governance for explanation of behaviour, the private behaviour and the related private mode of observation, as opposed to a public mode, would be assumed, which would necessitate the inclusion of both phenomenological experience level and intentional terms as part of explanation of behaviour. Intentional behaviourism does not need to be in a direct conflict with concepts of discriminative control and learning history in radical behaviourism – it is that but also more, as it involves setting variables such as discriminative stimuli, motivating operations, and implications they carry in terms of reinforcement. However, it would be worth to discuss whether the intentional explanations of behaviour could be considered *causal*. It is commonly understood by many philosophy of mind disciplines and is an underlying assumption of behavioural analysis as well that experiencing and attitudinizing can be recognised to serve as causal factors of behaviour. This can be closely related with the notion of *reasons* such as feelings, beliefs, or desires serving as causation of behaviour; and even though it is not entirely clear in what way specifically reasons cause behaviour and the mechanics and the process are not yet fully understood. Even so, the very existence of causal relationships should be undeniable: otherwise, if reasons did not serve as the cause for the behaviour to be the effect, and in the absence of the cause there would be absence of effect, then causal relationships would not need to exist. However, even if philosophically conceived mental constructs were a necessary prerequisite for the behaviour that follows afterwards to occur, not all thinking produces behaviour: there are myriads of thoughts that do not materialise in any manner, while other non-mental functions such as learning history are able to carry causal capacity, thus questioning this account as a plausible explanation of causal relations. In science, causality can be demonstrated with an experimental method within extensional context – when employed to study human behaviour however, the results, as they intermix with mental explanations of causality, require deliberate interpretation. As radical behaviourists accept the possibility to attribute causal effects to public events, and individuals can be assumed to be able to modify these publicly acquired rules, it should be possible to argue that private verbal behaviour carries the causal capacity, and the rules formulated in such a way are in fact intentionalistic in nature. Nevertheless, this does not suggest that intentionality is causal, but rather that behaviour would be predictable in terms of radical behaviourist contingencies when the rule-forming process agrees with the intensions ascribed. Even though it is not entirely clear whether contingencies may or may not be modifiable by the rule formations, it should be apparent that the intentionality is able to contribute another dimension to the explanation of behaviour.

2.6.3 Intentional behaviourism

Thinking and feeling are the personal subjective experiences that can be used to ascribe meaning to the observed experiences of others, thus attributing description to their behaviour in attempt to explain and predict it. In commonsense psychology, even though such relations cannot be proven, they can be taken as causation of behaviour nevertheless – something that is seen in intentional behaviourism as nothing more than a placeholder rather than a causal element without a proper validation. Intentional behaviourism is not concerned with ontology but rather considers thoughts and feelings as linguistic concepts that carry the capacity to serve as modules of behavioural explanation employing subjective experience: theoretical contributions can be developed from contrasting the sentences that use the intentionalistic language with those that do not. Radical behaviourism wholly relies on the use of extensional language and frameworks, and some behaviourists rely on neuroscience which is expected to provide a yet to be discovered physiological neural basis of behaviour. As such, intentional behaviourism does not propose to consider anything beyond the intentionality in terms only to identify the evolutionarily consistent neural functions without attributing causal properties to personal events, and therefore does not consider a neuro-physiological level of behaviour.

Intentional behaviourism is a philosophy of psychology that follows the original arguments proposed by Dennett in 1969 (reprinted in 2002) in his attempt to resolve the matter of intentionality within the analytical framework of conceptualisation, where it is claimed that it is necessary to describe a certain dimension of human behaviour with intensional language set against the extensional language of radical behaviourism. The use of intentional idioms to explain elements of behaviour on a personal level should be seen as adding intentional content in a systematic manner on another interpretative level entirely – thus contributing to and being consistent with the explanation offered by extensional sciences such as neurobiology. Hence, using intentional ascription would not be a simple matter of additional interpretation to the extensional description: explanation confined to extensional theoretical framework would be able to provide a descriptive and predictive scientific account of the behaviour in terms of structures and systems, whereas explanation available from the contribution of intentional theoretical framework would be able to offer the understanding of the actions and what the behaviour is accomplishing – in addition to the extensional explanation of how it is able to accomplish it. It is a matter of developing a comprehensive account of behaviour, and intentionality would be able to contribute substantially on the explanatory dimension in attempt to develop an all-inclusive explanatory account, while extensional framework of radical behaviourism has clearly been able to offer an extensive programme of predictive and behavioural controlled capability. For example, describing the continuity of human behaviour as a pattern with a certain goal or achievement in mind is no doubt an intentional in nature type of explanation – an explanatory method that radical behaviourists proposed (for example Rachlin, 1994) or perhaps unintentionally employed before, which is effectively the equivalent as the theoretical framework of intentional behaviourism.

2.7 Super-personal cognitive psychology

In this section, the link between empirical science of radical behaviourism and the philosophical framework of intentional behaviourism is discussed, and corresponding model of super-personal cognitive psychology is explained and discussed in some detail.

One important distinction that sets super-personal cognitive psychology aside and differentiates it from intentional behaviourism is that super-personal cognitive psychology is able to provide a decision process necessary to influence the observable behaviour for those aspects of theoretical framework that would benefit from the intentional explanation as identified by intentional behaviourism (Foxall, 2009). Thus, super-personal cognitive psychology makes it possible to specify the cognitive operations in a manner consistent with physiological and behavioural frameworks while exposing the process of decision-making and choice. Following on the premise of radical behaviourism that anticipated advances in physiology would provide the necessary contribution to the explanatory dimension – without a properly defined framework in place such as super-personal cognitive psychology or alike there will be no structure to facilitate the process of incorporating the explanatory dimension that may come from the said physiological advances, nor there would be a mechanism in place to direct the physiological research efforts towards the more plausible and likely to bring successful results areas of study. Moreover, the framework is also necessary to make it possible to recognise and verify these very advances in physiological research, and also confirm the potential results as fruitful once the research programme reaches its ultimate goal (Foxall, 2009).

It would be of some relevance to the discussion about super-personal cognitive psychology to contemplate in what manner exactly is intentionality able to contribute to the explanation of behaviour if it intentions are taken to be non-causal. The explanatory dimension that intentionality is able to provide comes after the causal relations have been accounted for with extensional methods – the causality thus determined provides enough evidence to support the explanation of behaviour for particular instances and events, yet it is the intentional expressions employing linguistic elements such as thoughts and feelings that provide the explanation for the sequences of events and continuous behaviour, the personal level of explanation, and the limits of radical behaviourism. Even though it has been suggested that super-personal cognitive psychology should be useful to explore the possibility to reconcile the temporal and spatial disconnection of dependent and independent variables to determine causal relations in continuity of behaviour, and to provide adequate account of the behaviour-environment relations; the purely linguistic non-ontological nature of incorporating intentionality into extensional framework of explanation could effectively suggest that these two frameworks would not be able to function as a uniform performance theory and rather would have to be discussed using two separate linguistic and scientific modes of explanation. To circumvent this limitation, the elements of sub-personal cognitive psychology would have to serve as the basis for causal theory and carry a sufficient account capable to supplement the extensional science employing intentionality in those areas where the limitations have been identified in radical behaviourism. This can be achieved by demonstrating the crucial role intentional and cognitive entities play as part of causal relations that explain behaviour by either including them as additional variables of experimental analysis directly, or a more probable scenario – developing proxy elements comprised of extensional variables to symbolise the emergent intentional and cognitive properties of behaviour. If this can be corroborated experimentally and indeed eventually be taken as a factual case, intentional and cognitive elements of super-personal cognitive psychology and intentional behaviourism could take a form of explanatory value comparable to that of extensional sciences, while functioning on a different level of explanation – otherwise be substituted by the eventual future performance theories developed within the field physiology or other. Thus, the explanation that the extensional account of radical behaviourism is able to offer is predictive in nature and demonstrates the causal relations between the variables determined with the process of experimental design, whereas the explanation that the intentional account of super-personal cognitive science and intentional behaviourism are able to offer is not necessarily compliant with rigorous scientific approach of radical behaviourism yet imperative to the all-inclusive explanatory account of behaviour (such as addressing personal events and continuity of behaviour) and therefore can be considered to be interpretative in nature. An intentional system therefore is an entity capable of using the intentional dimension on a personal level rather than something that can be predicted using the attributes and factors that comprise the intentional dimension.

2.8 Cognitive interpretation of behaviour

Having discussed the rigorous scientific method of radical behaviourism, and the potential benefits that intentional psychology could contribute to the understanding and explanation of behaviour, it is important to consider some areas of behaviour that are likely to remain for the time being out of scope for the deliberations presented here, and rather within the realm of cognitive explanation of behaviour.

Indeed, some of the frameworks presented above would be as useful with the more complex elements of cognitive explanation of behaviour as they are with the interpretative elements, and would provide the framework of reference to structure the consequent inquiry. In the same way, certain aspects of cognitive explanation could be designated as placeholders for future discoveries in neuropsychology and physiology. Until that time however, explanation of behaviour as it is considered here can be said to form three conceptual phases in terms of explanation, as follows.

First phase would be associated with explanation of behaviour as it is understood in the realm of radical behaviourism, which can be explained as a product of learned associations between a stimulus and a response, and reinforced or punished as a matter of consequence. This is something that was investigated in detail as part of earlier research programme (Greene, 2011) where Informational and Utilitarian Reinforcement were modelled as comprising elements of the input layer, showing significant predictive capacity as part of the connectionist model.

Second phase deals with interpretative elements of behaviour explanation, and is the focus of the research programme described here, which builds upon the findings of previous research investigation (Greene, 2011) and now proposes to consider Informational and Utilitarian Reinforcement as an emergent property which can be modelled within the hidden layers of connectionist networks to study Informational and Utilitarian Reinforcement as an emergent phenomenon which is formed as part of a model learning process and is subsequently used to develop understanding and explanation of behaviour. This will be described in detail in the following chapters.

Finally, the third phase of behaviour explanation would focus on cognitive elements of behaviour, which can often be inherently subjective and not yet understood on a sufficient level. Therefore, application of a scientific method and any kind of work that would rely on experiments would be inadequate at this level. Instead, explanation could come from the area of qualitative psychology that relies on phenomenology and *meaning-making* as part of contribution to the wider research programme that aims to understand and explain behaviour.

2.8.1 Phenomenology

In philosophy, phenomenology is a school of thought that studies the structure of experience and consciousness, attempting to establish necessary conditions for the objective study of certain topics which can typically be seen as subjective in nature, such as consciousness and the content of conscious experience – for example perceptions and emotions (Husserl, 1970, 2012). The structure and essential properties of experience are determined through a process of systematic reflection – approach that aims to be scientific, yet deemed qualitatively different from the method applied in clinical psychology or neuroscience. A number of assumptions that form the foundation of phenomenology would juxtapose it with the rigorous scientific methods of radical behaviourism – such as the rejection of the concept of objective research; or the preference to explore the personal conscious experience over the traditional scientific data analysis, to study human behaviour through a unique way it reflects the society of a person.

Phenomenology is generally considered anti-reductionist – even though varying degree of reductions are used in many of the methods employed to describe the underlying mechanisms of consciousness, the ultimate goal is to explain how the different aspects of reductions are constituted as an actual phenomenon as experienced by the person.

Intentionality is of course an important element of phenomenology, and stipulates that consciousness is always *about* something – the object of consciousness is referred to as *intentional object*, which doesn't necessarily need to be a perceptible object and rather could be anything at which consciousness is directed and of which it is conscious of, such as an idea or a memory for example. Intentional object can be constituted for consciousness in different ways (perception, memory, retention, etc.), and even though these different structures can be interpreted as different *intentionalities* that prescribe different ways of being about the intentional object, the object is constituted as the same intentional object throughout.

Another central element of phenomenology relevant to the present discussion in particular is the notion of *intersubjectivity*. Perhaps best explained referring to the concept of *empathy* as it is defined in the context of phenomenology, which requires a person focusing on the subjectivity of another person as part of intersubjective engagement, relying on the apperception built on the personal experiences. Therefore, one person is said to apply their own experience as their subjectivity to the experience of another person through apperception, which can be constituted as another subjectivity, thus facilitating recognition of another person's ideas, intentions, thoughts, etc. This experience of empathy is essential in the phenomenological account where intersubjectivity effectively constitutes objectivity: what is experienced subjectively is also intersubjectively available for other subjects. Thus, the notion of objectivity is not reduced to subjectivity, nor it is sufficient to constitute a relativist position, and instead a person is said to experience themselves as a subjectivity in an objective existence.

For an extended discussion on philosophy of phenomenology please see Husserl (1970, 2012).

In psychology, phenomenology is a study of subjective experience. Even though philosophical psychology of late 19th century principally relied upon introspection, deliberations concerning the mind based on such observations were largely criticised as speculation by emergent research movement that strived to uphold psychology to a more rigorous scientific approach – amongst them pioneers of radical behaviourism. The central philosophical issue is the problem of *qualia,* which is often referred to as a subjective conscious experience, where it is impossible to confirm whether experience of one person such as a feeling or an interpretation of meaning about an object is the same as that of another person.

To retort these criticisms, it is often claimed that phenomenological inquiry is a qualitative approach to study psychological subject matter that deals with the process of how meaning is construed and therefore interpretative in nature.

2.8.2 Interpretative phenomenological analysis

One of the approaches in phenomenological psychology notable for combining idiographic, psychological, and interpretative elements is *interpretative phenomenological analysis* (Husserl, 1970, 2012). Rooted in the phenomenological and hermeneutic theory, qualitative research approach strives to examine how a phenomenon is experienced by a certain person. Studies normally involve only a few participants – sometimes even just one person – whose experiences are closely examined using in-depth interview, diaries, or similar qualitative unstructured techniques that produce a detailed verbatim account for subsequent research analysis.

Due to the nature of the approach, participants tend to share relevant experiences in a particular context around the subject of investigation rather than being randomly sampled, thus allowing to scrutinise how phenomenon is understood from a shared perspective, and sometimes developing the research design further to include elements of longitudinal analysis by collecting multiple accounts over time. Hypothesis or theory testing is not normally part of the analysis, and interpretative phenomenological analysis would rather be employed with research questions that aim to understand and interpret a certain experience and explain how it was understood by a particular person. Instead, a reflective and often theory-developing account of hermeneutic inquiry is carried out with a focus on meaning-making, as researcher dissects in detail participant's key claims and codes them, offering interpretations of reoccurring theme patterns and their implications – thus attempting to interpret participant's interpretations and forming a situation of double hermeneutic.

As a result, the subject-focused approach that relates phenomena to experiences of some personal significance produces an idiographic analysis that interconnects phenomenological examination with interpretative elucidations, frequently illustrating the points by using participants' verbatim quotes and contextual commentary and details.

2.9 Summary

In this chapter, the consumer behaviour analysis is discussed in detail, outlining the framework of Behavioural Perspective Model. Radical behaviourism is juxtaposed with intentionality and cognitive psychology to explain and identify the underlying philosophical and methodological underpinnings of both theoretical frameworks. Intentional behaviourism is proposed as a possible extension of radical behaviourist framework to address the shortcomings and limitations of extensional sciences by employing the intentional linguistic elements in attempt to develop a better all-inclusive explanatory account of behaviour.

Once it is recognised that explaining behaviour is not only limited to predictive account of causal relations but also involves the interpretative dimension to develop an all-inclusive explanatory account of behaviour explanation, using intentionalistic linguistic elements and irreducibly inferential terms as proposed by intentional behaviourism should provide as a result an entirely different account to what is otherwise available from extensional science such as radical behaviourism. It is not a matter of integrating intentional terms into the behaviourist approach as it is already the case and intentionality is common in behaviourist explanation, but rather a matter of how these intentional terms should be integrated into the philosophy of psychology and extensional science. Intentional behaviourism is proposed as a competence theory of behaviour, specifying the mechanism to explain complex human behaviour by attributing interpretative intentional content in a systematic manner consistent with the rigorous scientific method behaviour analysis and extensional sciences. Super-personal cognitive psychology provides a structure and a form for the anticipated advances in the field of physiological research to be shaped and centred on the crucial questions of radical and intentional behaviourism such as the links of intentionality and cognition with the underlying neurophysiological processes, and a framework to identify and recognise these advances when the time comes.

3. Connectionism and Artificial Neural Networks

The philosophical underpinnings of *artificial neural networks* (NNs) have been explored by the researchers in the field of *artificial intelligence* (AI) at considerable length (Luger, 2005). To understand the theoretical and philosophical foundation of connectionist models, it is important to consider the historical developments of the conceptual design that the very fundamental structure of connectionist modelling is based on. To do so, the development of the field of AI will be reviewed briefly from its very inception and the influences it has had on the development of the NNs models and methodology over the years.

3.1 Why Connectionism?

For a number of years, research in the field of consumer behaviour relied on the traditional statistical methods to generate a substantial amount of empirical evidence to support the theoretical framework in. So why consider connectionist models at all?

Curry and Moutinho explore application of neural networks to study and model consumer behaviour, and offer a comprehensive discussion of theoretical and practical implications (Curry & Moutinho, 1993; Moutinho, Davies, & Curry, 1996). Authors deliberate application of expert systems as one possible alternative, but caution about limitations and potential overoptimistic notions in the field. Instead, artificial intelligence based application is suggested: artificial neural networks. A typical connectionist structure that incorporates a number of hidden layers brings certain advantages through a more sophisticated platform for modelling consumer behaviour, as hidden layers are able to distinguish key conceptual phenomena predisposed to indirect measurement. Another relevant to consumer behaviour feature is that connectionist models are trained: either through a supervised learning process where example connections of input and output pairs are fed into the model, or otherwise through relying on clustering methods in unsupervised learning. The ability to extrapolate patterns from training sample data offers a superior position to connectionist models against to rule-based arrangement commonly encountered in traditional systems, making neural networks particularly appropriate in tasks that involve notions of cognitive behaviour or pattern identification. This is discussed in further detail in the following sections.

3.1.1 The ultimate purpose of AI

The ultimate aim of AI research is to develop machines (or algorithms) to the level of performance and ability comparable to humans in tasks such as vision, natural language processing, learning, planning, reasoning, and other. Some of these tasks may seem simple enough and even intuitive if considered by a person new to the field. If the focus is shifted to machines however, it becomes readily apparent how immensely complicated these tasks could actually be. In fact, it is currently unclear the extent to which the most ambitious of these are achievable at all. Consider driving a car for example – many things need be performed simultaneously including at the very least visual recognition of the road and obstacles, geographical location and overall direction and the route with the final destination, traffic and participants (drivers, pedestrians, road workers, etc.), and the list goes on. This is not even mentioning the adjustments necessary for the time of the day, weather conditions, in-car distractions and overall condition of the driver, etc. This task poses a serious problem for a machine, and yet millions of people perform this with a seeming ease – many on the intuitive level even (Gallant, 1993).

Original approaches of AI to tackle a problem of this sort were mostly concerned with decomposing the overly complicated AI task into simpler ones in a systematic manner: the road lines recognition with the adjustments for the light and weather conditions, followed by three-dimensional figure recognition task combined with the geographical positioning and route planning, etc. It however becomes apparent that it is unfeasible to account for all possible circumstances of the task following this method. Alternative approach is to devise a heuristic trick rule such as following the centre line. This however is arguably even more fragile: what if there is an accident, or there is no centre line to speak of at all? Upon consideration of the complexities it could be tempting to declare these AI tasks unsolvable – with the exception that billions of biological creatures perform them on a regular basis with seeming ease (Gallant, 1993).

It became apparent that a novel approach was required that would be able to cope with high demands of AI tasks. It is then only natural to consider connectionist computational models that are inherently similar if only in structure to human information processing faculties. Since it is not entirely impossible that connectionist structures are even required for some of the AI tasks, it is no surprise that NNs gained such a widespread recognition in the field of AI.

This is not to say though that NNs are *the* computational modelling method for AI, as science should always strive to develop better and simpler methods (Gallant, 1993). The characteristics of NNs described in the following sections however provide a supportive evidence that at least some of the AI tasks can be advanced and explanations further developed.

3.1.2 Networks and symbolic systems

The traditional approach to cognitive science mostly evolved in such fields as cognitive psychology, psycholinguistics, neuropsychology, philosophy, and artificial intelligence because they share certain core assumptions of symbolic cognitive representation. Connectionism challenges these assumptions and offers a new promising approach to cognition via parallel distributed processing and neural networks methods in a number of ways.

The theoretical framework of connectionist networks is based on our knowledge of nervous system: the basic idea is that the neural networks comprise of simple elementary units with certain degree of activation, which are connected to other units thus making it possible to excite or inhibit other units in a dynamic way. Depending on the network design, initial input is spread through the network until the particular state of equilibrium is achieved – which in cognitive and decision-making tasks is in itself a solution to a predetermined problem (provided an appropriate interpretation is available). Dissimilar to the symbolic systems, the connectionism based adaptive network systems do not follow the rules based approach – which in itself is a rather primitive and simplistic design and will be further discussed in later paragraphs. On the contrary, the adaptive neural networks follow a rather different system design and focus on causal processing where units excite of inhibit each other and for the most part do not take into account stored symbols and their governing rule systems.

3.1.3 The foundations of connectionism

Initially connectionist models were developed following the basic functionality of human brain. Given the very limited knowledge of how the human brain actually works, neural networks are not intended to model the brain in its all complexity but rather simulate specific cognitive functioning in artificial systems that are able to exhibit some of the basic properties similar to those of neurons and synaptic connections in the brain. The first models designed to utilize the connectionist framework showed how the models consisting of a number of interconnected simple computational neurons could solve logical operations *and*, *or* and *not*. Furthermore, it was demonstrated that any process that could be performed using a finite number of these logical operations could be also performed by the connectionist network provided the necessary characteristics are met (for example the necessary memory capacity). Further advances came from converting the neurons from a binary to activated by a statistical pattern based on a number of input units, and as a result significantly increasing network reliability through parallel processing which is built as an inherent design feature – an early example of *distributed representation*. Following the string of research that revolved around the formal characteristics of behaviour exhibited by the connectionist models, the potential applications for carrying out cognitive tasks were examined (Pitts & McCulloch, 1947). One of the central and very frequently researched tasks in connectionist modelling is that of pattern recognition. Rosenblatt was one of the early researchers to advance the connectionist theory, introduced the continuous connection weights to replace the binary nature of neurons of Pitts and McCulloch models, and explored the networks where the excitations could be sent backwards, referring to these systems as *perceptrons*. He also devised methods to adjust the connection weights effectively establishing the procedures to train the network. Through the milestone *Perceptron Convergence Theorem*, it was demonstrated that if a set of connection weights capable of producing a correct response existed, it would be possible for the network to learn the correct response through a finite number of repetitions (Rosenblatt, 1958). The *perceptron*, being based on statistical patterns over a large number of units and accounting for noise and variations rather than logical principles, was established as a new type of information processing system that was closest in explaining the functionality of nervous system and capable of having original ideas. Thus, the feasibility for non-human cognitive system with connectionist networks has been established within the field of artificial intelligence.

Another notable early researcher was Selfridge (1958) who explored the pattern recognition capacity of connectionist models. His model *Pandemonium* was tasked with the recognition of handwritten letters. The model performed the analysis in parallel and processed the features of the letter through the levels. First level dealt with feature recognition task and carried the outcome to the next level that gathered the information on the particular features for each of the letters. The notable characteristic of the model outlined is that it is still capable to perform a reasonable assessment even if some of the features were unordinary or missing altogether (Selfridge, 1958).

In addition to pattern recognition tasks, it was acknowledged early on that connectionist networks might be useful in explaining the mechanics of memory and how the associations between the different patterns are stored. Hebb (1949, 2005) suggested that the synaptic connections between the neurons in the brain that are jointly active are strengthened – the string of research further developed by Taylor (1956). Taylor explored networks consisting of analog units with activations on a continuous range where the outcome suggested that networks are able to generate patterns similar to those of the units with which they are associated.

The significance of exploring the concept of cognition with connectionist networks should be assessed taking a number of wider research directions – as an overall course of inquiry which aims not only with modelling the brain directly, but also with the understanding the cognitive performance more generally, effectively establishing the foundations for later neural network and artificial intelligence research programmes. In the 1960s and 1970s, however the symbolic approach remained the predominant paradigm in cognitive science.

3.1.4 Symbolic models

As alluded to before, one of the origins of symbolic models comes from logic and philosophy – logical systems comprising rules for symbol manipulation with a clear outcome deliverable. Deductive logic aims to identify a set of rules that would make it possible to generate the true propositions, and a system of such rules is referred to as truth preserving. Therefore, it should be possible to develop a system of procedures capable of displaying cognitive behaviour if intelligence depended solely upon logical reason. This however is not the case, as it is often required of humans to make estimated predictions, which would fall within the domain of inductive logic that aims to develop formal rules that lead from propositions that are known to be true to those that are estimated to be true. Following this, intelligent cognitive process could be thought of as a logical manipulation of symbols, where symbols are regarded as ideas with rules to govern them. The definition of a symbol has now changed with the application of computers in modelling where symbols are stored in memory and are extracted and manipulated according to the computer programmes without any considerations for semantics. Alternative approach to interpret the semantics of the computational model offered by Newell and Simon (1981) views computer as a physical symbol system able of not only following the prescribed algorithm, but also more importantly capable of using heuristic methods as shown in the work on artificial intelligence (Simon, 1977).

Thus, artificial intelligence research has its origins in cognitive sciences and symbolic models, but since has notably deviated through its pursuit of the idea that computers are symbol-manipulating systems in a more general sense. Artificial intelligence models represent the closest simulation of human cognition and show competence in such tasks as playing chess. From this, the two suggested outcomes are possible: on the one hand, the human brain is the only true symbolic system and computer is merely a very capable calculator capable to execute complex algorithms; on the other hand, the computer is the true symbol manipulator and human cognitive faculties are carried out in a different manner that merely resemble those of the connectionist models.

3.1.5 Connectionist models

The publication of *Perceptrons* in 1969 (Minsky & Papert) in some way or another had an detrimental impact on the direction of research in artificial intelligence. The pessimistic predictions outlined in the book contributed to the research efforts being concentrated on symbolic models instead – the turn of events proven to be unfortunate when later discoveries showed the predictions of the book being inaccurate. In the course of the analysis of network models at the time a number of criticisms were demonstrated – namely the inability of the two-layer network to evaluate certain logical functions like *exclusive or* (A XOR B is defined where A is true and B is not, or B is true and A is not). It is necessary for a network to include additional *hidden* layers – which brought upon the additional problem, as at the time no training algorithms for multi-layered networks existed. As a result, many researchers viewed these criticisms as an indication of a larger issue and network models came to be identified with the associationism deemed inadequate for effective cognitive modelling.

Nevertheless, in the 1980s papers employing network models to simulate cognitive processes started to emerge in the published research yet again (for example J. A. Anderson & Hinton, 1981), and subsequently increased substantially. The network model renaissance was prompted by a number of factors, among which are some of the following.

First, the emergence of new powerful network modelling and training techniques and architectures, coupled with the advances in the mathematical descriptions of parallel systems directly applicable to modelling of cognitive processes. Second, the credibility of researchers newly converting to network models played a crucial role. Third, the structural resemblance of network models with the arrangement of the human nervous system, facilitated by the increased interest of cognitive researchers in neuroscience. The growing diversity and complexity of rule-based symbolic models resulted in the nostalgia for the parsimonious theoretical ground (much like what behaviourism was set out to offer previously). Finally, a growing number of researchers started to scrutinize the limitations of symbolic systems that revealed such weaknesses as inflexibility, unwarranted complexity, domain specificity, insufficient generalization, and scaling issues due to searches in large systems. These and other notions were able to deter some of the cognitive scientists from the symbolic models resulting in the increase in the increased appeal of network modelling, reaching a sizeable presence by the end of 1980s. In response, the symbolic models introduced a number of modifications to address the criticisms outlined above, such as using the rules on a smaller granularity and employing the selection and weighing criteria, and even started to incorporate some of the network features. Some of the key differences between the network and symbolic models remain however, including the ordered symbol sequences and consecutive operations of symbolic models that are not part of network model structure.

3.1.6 Argument for symbolic models

In the field of artificial intelligence, the cognitive models have relied predominantly upon the symbolic tradition for a number of decades. With the re-emergence of connectionism as an alternative approach to cognitive modelling, many proponents of symbolic tradition have developed a body of research to argue the inadequacy and limitations of connectionist approach.

Two most prominent critiques of connectionism raised by Fodor and Pylyshyn (1988), and Pinker and Prince (1988) are discussed in the following paragraphs.

3.1.6.1 Symbolic representation with constituent structure

Fodor and Pylyshyn (1988) argue that connectionist networks fail to fulfil the requirements of the *representationalist* system (intentional, semantic) without the capacity offered by a *symbolic* representation system, and therefore are inadequate for modelling cognitive processes. Symbolic representations have a language-like character, which Fodor refers to while hypothesizing on *language of thought (1975)*, and *combinatorial syntax and semantics* to govern the formation of molecular representations from the constituents. Composition and other syntactic rules can be applied irrespective of symbol semantics in a way that *syntactic engine* mirrors the *semantic engine* (Dennett, 1981) – something that connectionist systems are said to be lacking, as individual or groups of units cannot be developed into linguistic expressions that follow syntactic rules and composing simple representations into representations of higher complexity (Fodor & Pylyshyn, 1988). Thus, it is argued that only a system with symbolic representations and constituent structure is suitable for modelling cognitive processes - such as a language of thought that requires the following combinatorial syntactic and semantic features: (1) the capacity to produce and understand propositions from an infinite set, (2) the systematicity of thought that manifests in the intrinsic connections between the ability to comprehend one thought and other thoughts, and (3) the ability to make syntactic and semantic inferences. On that premise, the localist connectionist networks do not possess the necessary resources for cognition. In response, Smolensky's (1988a) criticism points out the oversimplification in their analysis of the networks by Fodor and Pylyshyn, as units in *distributed* connectionist systems are able to encode representational features and microfeatures, and therefore are more suitable as cognitive systems. In reply, Fodor and Pylyshyn argue that the ability of distributed networks to recognize the compositional microfeatures of an entity is not the same as the ability to identify one syntactic unit as part of a larger syntactic unit. Networks lack such syntactic relation, thus reducing connectionism to only an account of implementation of the symbolic representational system (on the nervous system level). In contrast to Fodor and Pylyshyn's account, Rumelhart and McClelland (1985a) distinguish between the level of information-processing account of behaviour and the level of abstract accounts. As such, it should be possible to suggest the multi-level account where abstract account is the subject of such disciplines as linguistics; connectionist and information-processing systems operate at the hierarchically lower level of analysis that is the subject of artificial intelligence and cognitive psychology; and neuro-physiological account at an even lower level (contrary to Fodor and Pylyshyn's only two levels where connectionism is assigned to the lower level).

Connectionism not only aspires to provide an adequate account of the phenomena that is successfully handled by rules; but also, without additional mechanisms, offers an elegant account of other phenomena as well.

3.1.6.2 Argument for rules

Pinker and Prince (1988) focus their critique around the children's language acquisition that necessitates the use of rules. In response to Rumelhart and McClelland's (1985b) connectionist model that simulates acquisition of English past tense and applies a uniform procedure for every case, Pinker and Prince develop and extensive analysis to determine whether it is a plausible model of human language acquisition. As a result, Rumelhart and McClelland's (1985b) past tense model was held to a much higher standard than it was intended to meet (that no other language acquisition model was able to meet either), ignoring the substantial development learning simulation that was attained with such a simple network architecture.

One line of criticism revolves around the type of decomposition of linguistic phenomenon in which rule-based and connectionism models differ, where Pinker and Prince (1988) argue that the mechanistic type of decomposition implemented in the connectionism model, in contrast to a more abstract decomposition employed by symbolic models, is inappropriate. Additionally, the ability of the network to analyse the phonological strings for patterning is attributed to the Wickelfeature (for an extended discussion please see Coltheart, Curtis, Atkins, & Haller, 1993) structure and not the network architecture. They also point out that Wickelphones (Coltheart et al., 1993) are limited to encoding phonetic information, disregarding syntactic, semantic, and morphological information important for past tense formation; thus limiting the ability of connectionist model to encoding the input-output patterns and not the abstract linguistic information. The fact that regular and irregular past tense forms are considered linguistically different in symbolic models, yet learned by the same mechanism in connectionist model, is also misguidedly seen as a shortfall.

Many of these concerns are addressed yet again referring to already discussed multiple levels of hierarchy assumed in connectionism: (1) neural processing, (2) information processing and (3) abstract level. The last point however is the central issue that Rumelhart and McClelland's (1985b) model strives to address – the ability to provide account of both regular and irregular forms with a single mechanism, irrespective to the linguistic prescription that necessitates different decomposition processes for the two forms. Further criticism refers to the already acknowledged limitations of the two-layer network. Hidden layers and back-propagation learning technique offer significant improvements of the network capacity (Rumelhart, Hinton, & Williams, 1985, 1988), and structured networks employ inter-network architectures (Touretzky & Hinton, 1988). In the end, Pinker and Prince (1988) reluctantly acknowledge that advances in network architecture and learning mechanisms may potentially enable the connectionist models to meet the criteria they specified, still unlikely to be able to provide more than a mere implementation of standard grammar.

3.1.7 Argument for connectionist models

In the next sections, three kinds of connectionist response are considered in response to the claims that prescribe the use of rules and symbolic representations as compulsory.

3.1.7.1 Approximationist approach

One connectionist view relies on the premise that symbolic models are abstract accounts of the phenomena and lose some level of detail in providing an efficient account of regularities, and therefore are *approximates* of the connectionist model account of cognitive performance that offers the highest level of detail (Smolensky, 1988a). Thus, in the case of language, it is the symbolic models that perform the task of approximation and connectionist models, once sufficiently developed, would be able to provide a full account of language. Rumelhart and McClelland (1985b) also promote this position, as in their view symbolic rule-based systems are too *brittle* and therefore unable to capture flexibility and subtlety of the cognitive process in its entirety, as cognitive behaviour is not governed by rules but is rather only approximately described by the rules at best. The behaviour is thought to be governed by a unified mechanism at a lower than rules level – sub-conceptual level (Smolensky, 1988b) or microstructure level (Rumelhart, 1975). Alternative approach would be to develop intricate rule-based systems that operate on a lower level and utilize soft constraints for the micro-rules (Holland, Holyoak, Nisbett, & Thagard, 1986)

Connectionist models described earlier that are able to attain considerable success in extracting the rule-like behaviour without the explicitly defined rules (for example past-tense acquisition model, Rumelhart & McClelland, 1985b) provide some preliminary evidence in support of the approximationist approach. The next step would be a model capable of syntactic processing without the reliance on rules at a level of performance comparable to a rule-based model in the very least. One such attempt is a network system for processing finite state grammar strings by Cleeremans, Servan-Schreiber, and McClelland (1989). One of the challenges for their network was the requirement to consider the preceding input while the current input is being processed – something a feed-forward network is not capable of attaining. To overcome such limitation, a novel architecture was devised by Cleeremans et al. (1989) as suggested by Elman (1989, 1990) – a *recurrent network* which, in addition to the regular feed-forward architecture, includes a subset of *context units*. These units do not receive external input, but rather receive activation from the hidden layer, making the previous interpretation of input by the hidden layer available during the current processing. As a result, the network was able to attain a high performance parameter and able to identify the preceding input in the string presented with appropriate training and network architecture. To better understand the learning process of the network, Cleeremans et al. (1989) used cluster analysis method to examine the information encoded by hidden units, which extracts activation pattern regularities that occur in hidden layers at specific input sequences to construct a tree cluster representation of similar patterns. Using just a few hidden units, the network was able to extract close approximations of abstract grammar rules - the outcome consistent with the symbolic perspective. Increasing the number of hidden units within the network architecture resulted in the learning patterns that develop a highly complex structure at different levels of cluster analysis, capturing not only the previous state of input but also the occurrence within the sequence. What Cleeremans et al. (1989) are able to show is that using the back-propagation learning, in the state of limited available resources (comprised of only a few hidden units), the networks resorts to extrapolation of abstract high-level rule-like patterns akin to that in symbolic models. When the network is not constricted however (more hidden units than minimally required for the task), it proceeds to extrapolate high-detail lower-level patterns that are more elaborate than the abstract rules in the traditional symbolic system.

3.1.7.2 Compatibilist approach

If the approximationist approach works from the bottom up, the compatibilist approach, on the contrary, works from the top down, and assumes at least some human explicit symbolic processing. Taking into account success of network models that do not use explicit rules, Touretzky and Hinton (1988) believe that it does not necessarily suggests the abandonment of explicit representations of rules in reasoning tasks entirely. Instead, embedded symbolic representation is implemented within a distributed subsymbolic connectionist architecture to achieve a powerful, parallel, fault resistant system of reasoning. As a result, instead of the usual training process where the network is left to its own devices to extract the patterns from the available data, in compatibilist approach the network is designed to implement the rules from the top down. In *production system*, (J. R. Anderson, 1981, 1983a), symbolic expressions are manipulated by explicit production rules. Touretzky and Hinton (1988) aimed to develop a connectionist implementation of *production system* – a system capable of using rules for symbol manipulation rather than approximationist network that generates approximate behaviour without reliance on symbols or rules. For exhaustive description of the *Distributed Connectionist Production System*, please see Touretzky and Hinton (1988).

3.1.7.3 Using external symbols for symbolic processing

Third alternative to approximationist and compatibilist approaches was proposed by Rumelhart, Smolensky, McClelland, and Hinton (1986) that revolves around the idea of networks developing the capacity to interpret and produce *external* to the network symbols. Natural language as a symbolic system fulfils a dual purpose as an internal and external tool – external symbolic formulations internalized through the *conscious rule interpreter*, which is a separate entity from the *intuitive processor* that operates on the inherent subconscious level (Smolensky, 1988b). Thus, it may be tempting to assume that humans need to function inherently as a rule processing system in order to operate as conscious rule interpreters. From the connectionist view, however, it may be possible to provide an alternative explanation of this account. Human developmental process occurs largely in social environments saturated with external symbols. As part of this developmental process, we learn the capacity to interact with external symbols by means of lower-level processes devoid (at least initially) of the symbol internalization ability, i.e. learning *how to use* external symbols. Even if it may appear that the use of symbols is eventually internalised to aid the reasoning faculties in mature adults, it is unclear in what way it is internalised exactly. Connectionist systems aim to identify and explain the causal relations of symbolic processing on a subsymbolic level – the step necessary to confirm that processing mechanisms at a higher symbolic level are in fact necessary. An alternative outcome could then suggest that connectionist pattern recognition ability may be sufficient to account for the symbol processing. Either way, the network approach to study external symbols may seem like a promising research direction.

The general idea here is dissimilar to compatibilist approach for the following reason: instead of developing a rule system, the network is trained to use a system that may contain the symbolic information such as rules. In such a sway, network is required to exhibit the usual behaviours such as pattern recognition working with the external symbols, and possibly benefiting from external storage function and other elements. Rumelhart, Smolensky, McClelland, and Hinton (1986) argue that the ability to simultaneously manipulate the environment and process the environment that is being created in the process using external symbols to solve difficult complex tasks by decomposing them into smaller simpler ones is the real and primary symbol processing function that humans are capable of performing. Thus, a system can learn to process and manipulate external symbols that are arranged according to some logical order. As it pertains to external symbol internalization, it is suggested that an internal mental representation of the external symbol environment is constructed, and the mental procedures operate on the internal representation instead. The output of mental model is used as input for the next mental procedure, and the output of that procedure used as input for the mental model, maintaining series of mental operations as a loop. Symbols are understood as patterns in the network, where the stable states of the network are symbols on a subsymbolic dynamic level.

The symbol manipulations are therefore treated as a learned capacity initially performed in the external environment, where symbols are the human artefacts that may be internalized similarly to nonsymbolic information.

3.1.8 The appeal of connectionist systems

One of the principal reasons why network models have generated an increasing interest within the cognitive science community is the demonstration of many properties that exhibit similar behaviour to human cognition which are not observed in symbolic models. There are a number of qualitative differences that set NNs apart from other AI approaches, namely the learning and representational abilities. Other distinguishing features worth noting are inherent parallelism and nonlinearity, and the ability to exhibit exceptional performance with noisy data (Gallant, 1993).

The next paragraphs provide some details on that aspect.

3.1.8.1 Natural plausibility

It should not be a surprise that network models are compatible with what is known about the human nervous system. After all, network models were initially designed to model the human nervous system and the brain: network activation and propagation is based on the elements that can be observed in nervous system. Other elements of connectionist network models do not resemble the biological elements of a natural nervous system, establishing a stronger position for the artificial and cognitive nature of network systems. Here, network models are examined as systems of the artificial tasked with modelling a cognitive process, thus the natural plausibility is not as useful as it would be otherwise in a discussion concerned with the neurophysiologic aspects of nervous system.

3.1.8.2 Soft constraints

A connection between the units in the network system, much like the rules in a symbolic system, constitutes a constraint between the two units: if the first unit is active, the second unit is constrained to be active as well (in an excitatory connection). Rules however are of deterministic nature, so if the antecedent for the rule is satisfied – it is to trigger the consequent action is sure to occur. Network connections serve similar purpose, but, unlike the rules, receive input from a multitude of other units and therefore represent the situation with multiple constraints. Thus, the best solution is determined by multiple constraints and therefore does not necessarily constitute an optimal solution for each of the individual constraints imposed – this is usually referred to as *soft constraints*.

Soft constraints seem to be better suited for cognition modelling in a number of tasks: decision-making is but one such case where a person is often confronted with multiple alternatives. In network models, soft constraints provide a natural way to account for competing alternatives without specifying the underlying rules that govern the competition, at the same time not limited by the constraints of linear models in symbolic systems. Another benefit that soft constraints are able to provide is the improved performance of the system while dealing with information previously not encountered. The implementation of soft constraint, which is an inherent characteristic of network models, provides the ability to overcome some of the difficulties of symbolic models, such as exceptions to the rules (particularly in psycholinguistics) or common mistakes for example. Soft constraints of connectionist models tend to override these limitations and account for both – the regular behaviour that can be governed by rules, and the exceptions – with a single mechanism, where different connections carry out different functions in alternative contexts. Thus, connectionist network models that employ soft constraints are able to overcome some of the limitations of the inflexible rule-based symbolic models.

3.1.8.3 Graceful degradation

Network systems exhibit a range of features similar to the functionality of a human brain when it reaches the limit of performance. Normally a very reliable system, while under overwhelming strain or physical damage, human brain begins to show less than optimal performance rather than crashing – some of the requests or parts of information are ignored, affecting performance according to the level of overload. Similar behaviour can be seen in connectionist networks: when elements of the system are destroyed, it is very rare indeed to observe the total loss of specific function, and instead manifested as a nonspecific gradient loss of functionality and increasing limitation of abilities – the effect generally referred to as *graceful degradation*.

Symbolic systems are not able to perform in such a manner. If a rule is eliminated, the system loses the functionality this rule is able to provide completely. Redundancy and error-checking mechanisms are able to cope with system damage to some extent – still failing to exhibit the full effect of graceful degradation nonetheless. Whereas destroying connections or even units in a connectionist network, on the contrary, does not significantly deteriorate the performance of the system overall. Deleting particular units would remove the locally encoded information, but deleting connections would result in graceful degradation. Subsequently, the network would still be capable to offer plausible solutions using the available information and learning rather than crashing.

System that employs distributed representation would even be able to exhibit only slightly warped behaviour if some of the units were disabled. With additional damage, system response accuracy would deteriorate further but would still be able to produce a response according to its current distorted pattern. Thus, connectionist systems possess an inherent ability of graceful degradation as a consequence of the network architecture.

3.1.8.4 Content addressable memory

The information that can be retrieved from memory using the cues that constitute parts of the memory itself is usually referred to as *content addressable memory*. Modelling this type of information system using the symbolic models poses a considerable difficulty. An example would be a filing system, where information is organized and stored according to some rule – usually a one-dimensional (chronographic for example), or two-dimensional (chronographic and alphabetic for example) at most. Accessing the information in any other way (not chronographic or alphabetic but rather performance based for instance) poses a considerable problem, as the information system was never designed for such a way and therefore would involve a serial search. Indexing systems could help to some extent, but this would require determining all possible paths for information retrieval on the onset, which could be unworkable.

Connectionist networks however, as discussed above, provide natural means to develop the content addressable memory system. It is even capable of dealing with certain inaccuracies with the cues and is able to provide the best alternative solutions if precise answer is impossible to determine due to inaccuracy. The effect is particularly apparent in distributed representation systems, where remembering a previous state effectively is no different from the process of making inferences and constructing a new state.

3.1.8.5 Learning from experience

Network models are capable of learning from experience by adjusting the connection weights, which will be discussed in detail in the following sections.

Symbolic models have been shown to be particularly suitable to represent learning that follows strict isolated rules such as instructions, whereas network models are particularly adept at large scale conceptual frameworks such as language acquisition. Further research is required to examine the usefulness of network models at acquiring instruction type memory.

3.2 The Neural Network architecture

Connectionist networks are adaptive systems comprising of simple computational units. Often containing thousands of interconnected units, unlike the traditional models with sequential processing, networks models are capable of displaying complex behaviour even with just a few units due to its parallel architecture.

3.2.1 Neural Network model structure

To illustrate connectionist processing, consider a model designed to simulate the functionality of a content-addressable memory system: the hypothetical dataset describes two groups of people with demographical and occupational attributes.

3.2.1.1 Network components

To have this data encoded, the connectionist network uses a number of distinct components. A set of (1) *computational units* is joined by (2) *connections*, and at certain times units examine its input and computes (3) *activation* as an output, which is passed to other units along the connections. Each connection carries a certain signed (4) *weight*, which determines whether the activations influence the receiving computational unit in a similar or opposite way according to the sign of the weight; and size of the weight determines the magnitude of the influence upon the receiving computational unit. Connections and weights are the imperative parameters of the model and determine the model behaviour.

3.2.1.1.1 Activations

Activation values for the units are determined by the equations. Initially set to a certain value, activations change once the simulation is run and are adjusted accordingly in response to the effects of external input, propagation of activations exhibited by other units in the system, and decay over time. Only the input layer could be affected by the external input – units in the hidden layer (each unit representing the group member for example) are only affected by the propagation of activation from other units and decay over time.

3.2.1.1.2 Connections

Weighted connections transfer the activations between the units. In the hypothetical dataset described above, the units for each of the group members with each of their attributes are excitatory connections. As a result, property unit propagates activation to the unit that represents a group member that possesses such attribute. To those units that represent a group members that do not possess such attribute the connection is inhibitory, as are the connections between the mutually exclusive attributes. Thus, if certain age group is activated, all the units that represent other age categories will become less active (due to inhibitory activations between the mutually exclusive properties) and units that represent group members that are of the appropriate age will be activated (excited) and those that are of a different age will not (inhibited).

3.2.1.2 Dynamics of the network

To illustrate the functionality of the network, consider a sample task of memory retrieval. The external input is supplied to one or several input units of the network. As a result, the excitatory connections will transfer the activation to the units associated with the input unit supplied with external activation and inhibitory connections will decrease the activations with the units that are not associates with the externally activated unit. Thus, if the unit representing group member's name is activated externally the activation will be carried to all the units that represent group member's characteristics, and will decrease activation with all the other units that represent other group members and characteristics other that the associated with the particular group member that is being externally activated. These effects will continue to reverberate throughout the network across numerous cycles until the system achieves the state of equilibrium where additional cycles no longer improve the performance. All the attributes associated with the externally activated group member will now have high activation values, thus recovering group member's characteristics from the system.

A more practical query may be the reverse to the one described above is the task of content-addressable memory, where the group member's name is retrieved through activating the characteristics. Simultaneously activating a number of characteristics will activate the hidden unit that represents the group member, which in turn will activate the group member's name unit. Moreover, it will activate all the units in the hidden layer that represent the group members that have excitatory connections with externally activated attributes, thus activating all the group members' name units to a varying degree – behaviour of the system that offers high level of generalization.

The process outline above is able to produce more subtle effects similar to human performance tasks of categorization and prototype formation. The network is able to obtain category instances with the external input to one of the category units, highlighting all group member units. During this process, all group members are characterized as to how well they represent the category. As category unit is activated, the activation is propagated to all member units, which in turn propagate the activation to all their characteristic units. Therefore, the most common group member characteristics become activated the most, sending their activation back to the individual member units thus forming the prototype.

Another example shows the practical example of utilizing regularities. If multiple characteristics units are activated externally, they will immediately propagate activation to those individual member units that share the activated externally characteristics, thus changing the activation for those members. In turn, the member units will propagate activation to other properties that are characteristic to the activated member units. This then will activate additional member units that do not possess the initial characteristics activated externally, but would identify the other group members who are most likely to show the best fit with the primary group. This network behaviour effectively allows making inferences from known characteristics to other characteristics.

3.2.2 Neural Network architecture design features

Illustrated in the previous network architecture may very well be very fitting to the memory retrieval tasks and provides attractive network behaviour associated with the task. This design however is not particularly suitable for most other applications. Other connectionist architectures may be distinguished with regard to the four features: (1) the way units are interconnected, (2) unit activation mechanisms, (3) learning procedures that alter the connections, and (4) the semantic interpretations of the system.

3.2.2.1 Connectivity pattern

Connectionist networks may be divided into two major classes in respect to the way the units are connected: (1) *feedforward networks* that have one-way connections where the activation goes from the input layer to the output layer and activation is forward propagated, and (2) *interactive networks* that have two-way connections where dynamically changing activations reverberate between the units in the network over many cycles.

3.2.2.1.1 Feedforward networks

Units in a feedforward network are arranged into distinct layers – the simplest two-layer configuration consists of only input and output layers. All input units are connected to all the output units and once the connection weights are configured appropriately, the network is able to produce a suitable response to the input pattern with a distinctive output pattern and for that reason is sometimes referred to as *pattern associator*. Pattern associator could be used as a classification device where inputs are sorted into few output categories. The limited architecture of two-layer network however is insufficient to address certain problems such as XOR (exclusive or) function. To accommodate such limitation, it is necessary to introduce the hidden layer into the network configuration. Situated between the input and output layers, hidden layer modifies the information processing and provide considerable additional functionality to multi-layered feedforward networks. A number of modifications are possible in the multi-layered networks. One such modification connects units not only to the next layer, but also to the units in the layers beyond the next one

– so in a three-layer network input units would not only connect to the units in the hidden layer, but also directly to the units in the output layer in addition to the connections between the hidden layer and the output layer. Another modification establishes the *recurrent network*, where the system receives the input in a sequential manner and response is altered according to the information of previous steps of the sequence. The pattern achieved on a higher layer is fed back into the lower layer and serves as a form of input.

3.2.2.1.2 Interactive networks

In contrast to feedforward networks, interactive networks include at least some number of two-way connections and input is processed across multiple cycles. If processing units are organized into layers, the processing could go forwards and backwards.

3.2.2.2 Activation rules

Another difference characteristic to network models in addition to the pattern of connectivity is the rules that govern the unit activation values. Activations values could be grouped into classes that include (1) discrete activations that typically involve a binary value (for instance 0 and 1) or (2) continuous activations, either bounded (a range of -1 o +1) or unbounded. Activation rules specify the calculation for the level of activation for each of the units. The following paragraphs will discuss the rules in detail.

3.2.2.2.1 Activation rules in feedforward networks

The input is composed of the two components: the external input and the effect of activity in other connected units. In a two-layer feedforward network, input layer units are dedicated to receive external input and take the input pattern value as activation and therefore does not require an activation rule. On the contrary, the output layer units are dedicated to receive activation from other units in the network. The output from the input layer units is sent to the output layer units, where it is multiplied by the connection weight – summative output from all the input layer units provides the total input for each of the output layer unit. Additionally, a bias could be introduced to regulate the responsiveness of each output unit as an additional input for the output layer units that are unaffected by the dynamics of the network: low value will result in conservative response of the output unit whereas high value will do the opposite. Activation for each of the units is then determined by applying the activation rule using the summative output from the input layer units. If linear activation rule is used, activation equals the summative output of input layer units – provided certain constraints are met. Introduction of hidden layers provides the additional power necessary to violate the constraints, making it necessary to change the activation rule accordingly to a nonlinear function – logistic for instance.

These rules could be adapted to be used with discrete rather than continuous activation values (for networks that use binary units). With the linear activation rule, the continuous output of the unit is compared with the specified threshold value unit and depending on the result; the continuous output of the unit is converted to binary (either 0 or 1). It is possible to use the threshold units within the hidden and output layers of feedforward networks, as well as in interactive networks. If the logistic activation rule used in a feedforward network with binary units, the binary activation takes the probabilistic form where the equation determines the relative frequency of the discrete result.

3.2.2.2.2 Activation rules in interactive networks

In addition to the equations used in feedforward networks, interactive networks necessitate the parameter for time (t) or cycle(c), as activations are updated numerous times for each of the units in response to the particular input. Unit activations may be updated once per cycle if *synchronous update* procedures are employed, or each unit is updated separately according to some random determinant in the case of *asynchronous update* procedure being employed (which helps with preventing the unstable oscillations of the network). Another difference is that each update requires a separate application of activation rule to be performed; unlike to feed-forward networks where there is only one forward wave of activation changes and once for each of the units. Hopfield nets for instance (Hopfield, 1982) comprise of linear threshold units where on each input unit acquires an activation of *1* if the input is above the threshold – otherwise activation is set to 0. Asynchronous update procedure is then employed for units to determine a random time to update activation according to the state of input of the network at the time of an update until none of the units would receive an update that would lead to a change of the activation. It is then the network is said to achieve the state of *equilibrium*, which constitutes the network's identification of the initial input (if the network indeed settles into the equilibrium and does not rather oscillate between multiple configurations).

An imperative role in acceptance of the model played the fact that Hopfield (1982) demonstrated the measure of the network state (*energy, E*), showing the analogy between the network ability to achieve equilibrium with that of the physical system - state of lowest energy in thermodynamic system. The *E* measure could be adapted to show the *goodness of fit* (*G*) of the network's end state of equilibrium (Rumelhart et al., 1986). Hopfield nets are increasingly useful in a number of optimization applications where the connections represent the constraints for possible configurations of network equilibrium (a solution to supplied input).

One of the difficulties Hopfield nets demonstrated is that the network can settle into local minima – the stable state where different parts of the network settle into incompatible configurations and as a result, the network is not able to achieve the overall state of lowed possible *E* value. To reduce such network tendency, Hopfield net has been adapted by Hinton and Sejnowski in their Boltzmann machine (1985; 1984). The difference with the Hopfield net is that it employs a stochastic activation function rather than a deterministic one – essentially, it is a probabilistic version of logistic function discussed in paragraph on activation rules in feedforward networks above.

Anderson's spreading activation models (1981; 1983b) that utilize negative exponential function of current activation to achieve nonlinearity in semantic networks, using decay function for interactive processing. Used in a service of a production system, the networks allows parallel processing within the system architecture by tolerating a number of similar to some degree active processes to run simultaneously in competition (similar to notion of *soft constraints*). Spreading activation models could be distinguished from the network models on the following criteria.

First, spreading activation models maintain the structure of control, whereas connectionist networks are dedicated to retain no control over cognition modelling other than internal decentralized local control of the network. Second, certain types of distributed representation is emphasized in connectionism (McClelland, Rumelhart, & Group, 1986). Third, propagation equations contain differences.

To summarize, all the different types of networks have certain common elements. New activation of a unit is dependent upon net input received from other units, which in turn is determined by the connection weights. Connectionist networks usually have the ability to alter their connection weights in adaptive manner through a process that is often referred to as learning. Different learning procedures are discussed in the following sections.

3.2.2.3 Learning procedures

In connectionism, the process of learning signifies the ability of the network to modify connection weights between the units. The weights determine in some measure the end state a network could reach as a result of the processing, and therefore transform the network characteristics. It is the goal of a learning procedure to define a basic procedure for the network capable to achieve the desired output without the external control system – that is a local system of weight change control. Readily available inputs of each of the units include the current value of the weight itself and the activations of the units to which it is connected.

Donald Hebb proposed an idea that suggests that learning occurs in the nervous system through strengthening of the connections between the neurons whenever they fire simultaneously. Based on this proposal, one of the simple learning procedures in connectionism specifies the weight of the connection between the two units is increased (or decreased) in proportion to the product of their activations – the *Hebbian earning rule*. Consequently, when both units have the same sign the weight is increased proportionately to the product of their activations or decreased in the same way when the signs are different. Although capable of producing impressive results, Hebbian rule is presents some serious limitations. A number of differing learning procedures will be discussed below – all however based on the same principles that specify learning as procedure of changing connection weights employing only the information available locally.

3.2.2.4 Semantics of connectionist systems

If a connectionist network is to simulate human behaviour or cognitive performance, one must consider the representation of the concepts of that domain in the network. It is possible to either designate each unit to a particular concept in *localist* networks; or designate multiple units to represent the concept, as is the case in *distributed* networks.

3.2.2.4.1 Localist networks

In localist networks, each concept is represented by a designated unit. One of the obvious advantages that this offers is the considerable ease with which researchers could monitor the network performance in terms of the domain and objects studied. This however carries a possible caveat in a way that it is easy to forget that the appointed to each unit concept only carries meaning to the researcher and not the network itself. It is necessary to rely on external system employed for semantic interpretation, which could limit the performance of the network by the design of the architecture. Despite this, it is may be even more difficult to design a distributed network, and localist networks may be preferable in a range of tasks. Concepts are represented by units, and constraints between those concepts are represented by the connections – positive connection emphasizes the condition of the network where both units have the same activation, whereas negative connection emphasizes the preference towards the opposite activations. The state that the network achieves at its global minimum is the state that best satisfies the soft constraints.

3.2.2.4.2 Distributed networks

In the distributed networks the situation is quite different, as the concept is represented by an activation pattern across a number of units rather than a single unit representing the concept. One way to design a distributed representation of a concept in a network is through featural analysis of the concept (method often employed in the field of psychology) to encode across the appropriate units. The separate features derived such an analysis are usually encoded as individual units in the network architecture – so are a localist representations of the features that form a distributed representation of the concept. On way to obtain a workable featural analysis of a concept is to rely on established theoretical framework, as often connectionists are not concerned so much with the features of the concept but rather with how the network utilizes the distributed representations. Another way is to allow the network to perform the analysis. During the learning process when only input and output parameters are identified by the researcher, hidden units of a multi-layered network will develop sensitivity to certain features of the concept. Network learning is effectively an intricate feature extraction mechanism, and usually networks do not attain the apparent from the input localist solutions but rather each of the hidden units develops sensitivity to a complex and understated regularity commonly referred to as *microfeature*. This, hidden layers are able to provide distributed coding of input.

Such distributed representation design carries additional benefits to the model architecture. Once the concept is distributed pattern across units, the processing capacity of the network is also distributed and therefore is capable to compensate for the missing, partial, or even inaccurate data through processing in other units. Thus, the system is more resilient to failure. Moreover, the system is able to learn new information or provide a response to previously unseen input. This network capacity is akin to human process of making *generalizations* – ability to infer some of the unknown properties of the entity based on the known properties – and therefore may be particularly suitable in such tasks (McClelland et al., 1986).

Another technique employed in distributed networks is *coarse coding*, and is discussed elsewhere (Touretzky & Hinton, 1988).

3.3 Machine learning

Machine learning broadly refers to ability of a model to improve its performance based upon input information. It is generally considered that research on machine learning presents the highest potential to eventually develop models able to perform complicated AI tasks, as algorithms that learn from training and experience are superior to those based on a subset of contingency rules developed by human scientists. Machine learning may be divided into supervised and unsupervised learning.

Supervised learning is a learning algorithm that analyses a training data (i.e. labelled data: pairs of input and output values) to produce an inferred or a regression function able to predict the correct output for any input. It is required for the learning algorithm to make certain generalizations from the training data that could be used to analyse previously unseen data – a process that is analogous to concept learning in human and animal psychology. Unsupervised learning refers to the machine learning problem aimed to determine the underlying structure of unlabelled data. In unlabelled data there is no error signal to evaluate possible solution, and therefore relies on techniques such as clustering that examine the core features of the data – the self-organizing map (Kohonen, 1990, 1998) is one such algorithm often used in NNs models (for a strategic marketing application, see Curry, Davies, Phillips, Evans, & Moutinho, 2001).

The capacity of Neural Network Models to learn is one of the features most notable to researchers. This complex question requires particular attention, and learning algorithms, both connectionist and other, are discussed here. Certain philosophical issues concerning the connectionist learning are also addressed.

3.3.1 Traditional approaches

Following the two distinctive philosophical approaches to learning, the theoretical findings in disciplines such as psychology and linguistics (and others) have customarily been divided to follow one of the two major intellectual traditions: the *empiricism* or the *rationalism*.

3.3.1.1 Empiricism

The philosophical *empiricism* (largely based on the work of Bacon, Locke, Berkley, and Hume) refuses the excessive reliance on established principles of reasoning, and views the sensory experiences as primary requirement for the acquisition of knowledge. Certain integral elements of the theoretical framework however pose a particular interest. *Associationism*, for instance, describes the sensory processes to result in simple ideas, which in turn are composed into complex ideas through the spatial contiguity that produces the association. Temporal contiguity is viewed as an integral part to the concept of causation, as idea of a cause would elicit the associated idea of effect. Simple ideas that are sensations are composed into complex ideas through simple additive mechanisms, and therefore are sufficient to predict the properties of complex ideas.

Behaviourist models employed a kind of associationism in such a way that the entities involved in the association were limited to those available for the observation – the environmental events and the behavioural responses. During the period dominated by the behaviourist theories, learning was one of the central research domains. In behaviourism, learning could be defined operationally as a change in response frequency. Researchers investigated different ways of arranging the environment and employed mathematical modelling to establish the theory, for the large part deliberately ignoring the internal processes and mechanisms of the system. This and some other limitations made behaviourists susceptive to the emergent information processing theories.

3.3.1.2 Rationalism

The other major philosophical tradition that influenced the development of cognitive sciences is *rationalism* (represented by Decartes, Spinoza, and Liebniz). Contrary to empiricism, ideas in rationalism are not restricted to experiences but rather are innate: what is important is how these ideas are used in reasoning. In psycholinguistics, Chomsky (1957, 1968) introduced the concept of innate

Universal Grammar, arguing that the amount of information that children receive in their early years would not be sufficient to develop the grammar rules of a child (poverty of the stimulus argument, in review of B. F. Skinner's Verbal Behaviour, 1959). Following the Chomskian tradition, the child is said to be born with a set of default parameters that could be reset according to experience.

Neither empiricists nor rationalist frameworks were able to provide a convincing account of the mechanisms of the language acquisition process.

3.3.1.3 Contemporary cognitive science

Unlike empiricists and rationalists, cognitive psychologists and artificial intelligence researchers have for the most part seemed to ignore the learning process until recently, addressing other areas where immediate result could be made using symbolic rule-based models (information representation, memory systems, etc.). The rise of connectionist approaches to learning in the 1980s generated and increased interest to learning. Research area known as *machine learning* emerged within the artificial intelligence framework that focuses on developing strategies for the machines to learn from experience. In rule-based systems, learning strategies focus on addition or modification of the rules – a rather challenging task, as modifying the rules to accommodate certain circumstances may result in a drastic deteriorating effect elsewhere. Another critique is that adding and modifying rules is arguably too rudimentary of a mechanism to capture the learning process.

Therefore, in the 1990s, empiricists continue to develop increasingly sophisticated methods to modify the symbolic rules; rationalists offer novel interpretations of adjustments to innate grammar system in language acquisition; and connectionists develop new algorithms that are able to provide the sub-symbolic network architectures of the learning process.

3.3.2 Neural Networks approach

Learning in connectionist models is a process of adjusting connection weights between the units, which would have an effect on subsequent processing of the input by the network. While the network is trained, the activations and weights change on each trial – subsequently after the network training is complete, the network is tested to examine the effect of the input on the activations only. Both weight and activation changes are determined based on the local information available immediately to each unit: remote units in the network are affected by spreading local propagation. A number of learning procedures in connectionist networks have been developed, and employ either *supervised* (classified according to specified input-output) or *unsupervised learning* (no feedback on input-output provided).

3.3.2.1 Two-layer feedforward network learning procedures

The objective of the learning procedure is to determine the weights to allow the appropriate response of the network to a number of cases. Each case is comprised of the input layer pattern of activations and output layer of pattern activations that form the *n*-dimensional vectors, where *n* is a number of units.

The learning ability is normally assessed during the training and testing stages. During the training, the weights are successively modified according to the constraints of the set of cases. Depending on the learning procedure used and the difficulty of the set of cases, it may require a large number of epochs for a network to achieve the desired state and, once the level of acceptable performance is attained, should be able to respond to the input patterns with the appropriate output pattern. Some particular learning procedures are the *Hebbian* and *delta* rules.

3.3.2.1.1 The Hebbian rule

The two-layer feedforward network that uses the linear activation and Hebbian rule to a set of input-output cases forms a learning system referred to as *linear associator*. During training input and corresponding output pattern is presented, and Hebbian rule is used to adjust the connection weights: the activations of the two connected units are multiplied with the learning rate. Consequently, if the two unit activations are both positive or negative, the connection weight will be adjusted by the amount specified; if one unit is positive and another is negative, the connection weight will decrease according to the negative equation value. During the testing stage, only the input patterns are presented. Hebbian rule offers good results as long as the input patterns are not correlated – a requirement that imposes a substantial limitation.

3.3.2.1.2 The least mean square rule

Similar to Hebbian rule in a way that it considers the input and relevant output unit for a change in weight, but substantially more powerful, is the *least mean square* (LMS) learning rule. During the training stages, the LMS rule generates the *actual output pattern* using the input pattern and compares it to the *desired output pattern*, and changes the weight accordingly to minimize the discrepancy for each of the units. Thus, the rule effectively is an error correction procedure.

Once the discrepancy for each of the units is calculated, they are squared and added together to compute the *pattern sum of squares (pss)* value. By summing all the pattern sum of squares values a *total sum of squares (tss)* value is obtained, which indicates the level of potential improvement still obtainable until the perfect performance is attained on the whole set of input-output cases. Thus, the essential principle is to change the connection weights as to minimize the total error, and LMS rule need not be restricted to the uncorrelated only sets of input patterns.

When the two rules are contrasted, the difference that gives the LMS method substantial increase in performance over the Hebbian rule is that the LMS method is able to utilize the discrepancy between the actual and desired output to change the connection weights during the training stages; whereas Hebbian rule is only able to use it for evaluation purposes during the test stages. Even though the two-layer network could be quite powerful given that certain conditions are met (linearly independent set – i.e. none of the input patterns are a linear combination of other patterns): it is capable of learning the *inclusive or*

(OR) function. It is however not capable of learning the *exclusive or* (XOR) function as there is no set of weighs capable of generating the correct output. The network will however aim to minimize the *tss* and will learn to generalize to new input patterns based on their similarity to the input patterns of the training stage. As a result, the network will do a good job in identifying the acceptable output pattern.

In the 1980s, the new powerful training algorithms for training hidden units such as *back-propagation* finally allowed to overcome the linear reparability constraint, leading to major breakthrough of connectionism.

3.3.2.2 Back-propagation learning procedure in multi-layered networks

By introducing the hidden layer between the input and the output, the information flow becomes increasingly more sophisticated. This allows the network to process intermediate results obtained from the input activations that are then used in the output. The network architecture becomes substantially more complex as well, as now there are a number of ways available to the researcher pertaining to the network design that need to be addressed. Number of hidden layers and hidden units in each layer is one such question, and researcher may choose to perform some exploratory analyses to determine the optimal model size. With multiple layers present in the network structure, it is necessary to consider the extent of interconnectivity between the layers in the network, as now not only the successive connections are possible (input-hidden-output), but also additional connections (input-output in addition to input-hidden-output). In addition, the more intricate network architecture requires a more sophisticated nonlinear activation rule. Finally, the learning procedure capable of handling the now available hidden units is essential, such as modified LMS procedure where the activations propagate forward and then error and weight adjustments propagate back through the network (*back-propagation*).

Despite the questions of the design, the actual network response is developed through the learning process by adjusting the connection weights and not determined by the researcher. An example of such network is NETtalk model (Sejnowski & Rosenberg, 1987) that was tasked with reading English. By supplying the continuous speech corpus of 1,024 words with desired output consistent with the phonetic speech, network was able to achieve 80 percent accuracy after 10,000 training trials, and 95 percent after 50,000 words presented to the model, with 78 percent accuracy on a previously unseen text. The voice synthesizer was actually able to produce recognizable speech using the model output. Further analysis exposed the functional features of hidden layers in the network being relevant to theoretically appropriate elements of language.

Some of the shortcomings of the back-propagation learning procedure that have been identified are concerned with a rather high computational demand and the fact that the network may take a long time to learn; as well as inability to distinctively attribute back-propagation to any known biological process. If viewed on a psychological level of analysis however, back-propagation is a mechanism that allows a multi-layered network to achieve *gradient descent,* i.e. learning by reducing the output error.

3.3.2.3 Boltzmann learning procedure

Interactive networks (such as Boltzmann machines) have their own distinct architecture different to that of the feedforward networks: each input pattern triggers numerous cycles of activation processing across the network, which maintain the interaction across the cycles until the network settles into the state of thermal equilibrium. Unlike the learning epochs in feedforward networks, the cycles do not involve the modification of weights but rather computation of activations, and respond to a single input pattern. One issue that may arise is the tendency of the network to settle into local minima – a stable state that nevertheless does not satisfy the constraints in the best possible way (addressed with the help of the simulated annealing technique that involves the gradual decreasing of the designated parameter in a stochastic function).

Boltzmann machines can be trained using a learning technique conceptually similar to that of a back-propagation (McClelland et al., 1986). While in training mode during stage one, the input and output units are fixed and other unit activations are updated in a random order using the stochastic equation with the simulated annealing until the network reaches the thermal equilibrium. Each of the input-output cases is then processed and simultaneous activation time is recorded as an expected probability of unit activation. In stage two, essentially the same process is carried out but only the input units are fixed this time and the output is determined by the network. The variation of the probability obtained in the two stages determines the connection weight change which will result in minimization of output units error (given the learning rate parameter is set sufficiently slow). Thus, the underlying reasoning is similar to that of the LMS procedure where discrepancies between the desired and actual output direct the adjustment of connection weights. One of the drawbacks of the procedure is the slow learning rate due to the time required for the network to settle into the equilibrium for each input pattern.

3.3.2.4 Competitive learning

Another type of learning procedure is competitive learning that is a variation of unsupervised learning, where a network is presented with input patterns and is tasked to identify regularities to allow grouping the patterns into clusters of similar patterns with no feedback on the correctness of the procedure. The simplest architecture would have a fully connected input and output layers, and the number of units in the output layer specifies the number of clusters for a network to identify, and the activation rule is set to ensure only one unit is chosen inhibiting the other units at the same time. Learning rule reallocates the weight of the chosen unit in a way that increases the connection weights with the active input units and decreases the weights with the inactive ones keeping the total weight constant.

Inclusion of more than one set of competing output units (possibly with a different number of units specifying a different number of clusters) would increase the complexity of the network behaviour. Multiple layers would also result in behaviour that is more complex and allow the network to identify higher order regularities.

3.3.2.5 Reinforcement learning

In reinforcement learning, the network is given the information on whether or not the output pattern was close to the desired pattern without supplying the actual desired pattern – therefore only the global performance indicators are used to adjust the weights. The procedure may not seem as advanced as back-propagation; nevertheless, it does satisfy the basic notion central to behaviourism and modifying behaviour through reinforcement. Essentially the network performs a large number of trials with varying weight combinations recording the global reinforcement delivered with each trial. The weight combinations that deliver higher reinforcement gradually become recognized, resulting in increased consequent trial frequency that leads to the weights that highest global reinforcement.

Reinforcement learning is much simpler of a procedure compared with the back-propagation since the error calculations for each of the weights are omitted; but may take a long time to produce the result and does not scale very well. It does however offer a substantial theoretical benefit of relating the connectionist method to the field of traditional learning theory.

3.3.3 Difficulties with machine learning

Some of the criticisms of connectionist models of learning are discussed in the following paragraphs.

3.3.3.1 Associationism

One of the critiques of connectionism is that it is essentially a return to associationism, which would have an adverse effect on the progress of cognitive science and the advances made by the symbol manipulation systems. It is not however merely a return to associationism, but is rather based on the core principle of associationism that suggests that contiguities produce connections. Connectionism employs the powerful idea and develops it with unmatched sophistication with such mechanisms and concepts as distributed representation, hidden units capable of capturing microfeatures, back –propagation procedures, supervised learning with error reduction function and other. A simplified connectionist network that uses Hebbian learning rule is most similar to classical associationism where simple units were ideas and, based on the contiguity, associations are increased or decreased between these ideas. The more sophisticated multi-level connectionist networks could attain deeper level of associationism as hidden units can decompose ideas into microfeatures and propagate their activity within the network to achieve contiguity in a different manner.

Rule-like systems could also be modelled with connectionist architecture on a micro level. Therefore, connectionism provides the mechanism that can operate on a fine level of detail and using the cognitivist high-level description of the cognitive process. Moreover, connectionist models of learning offer a novel approach to the process of concept and cognitive skill acquisition. Ability to provide plausible explanatory models of rule-like behaviour and offer powerful learning mechanisms may play an imperative role in cognitive science through the integration of associationism and cognitivism that may carry broad implications for the field.

3.3.3.2 Poverty of stimulus

In Chomsky's criticism of Skinner's verbal behaviour and language acquisition, the central nativist argument revolved around inability for a child to learn the language from available experience – the poverty of stimulus. The argument however should not be construed in a form whether anything in the organism is innate or exist prior to the sensory experiences, but rather what is innate since any kind of learning presupposes some sort of structure to be present within which the learning could occur. In the case of symbolic approach for example, at least some of the initial functionality in symbol manipulation could be considered innate.

In connectionism, nativism is rarely regarded as an issue – possibly due to the fact that connectionism has the roots in associationism. Something else to consider is that most challenging connectionist problems are statistical and computational and deal with the science of artificial in one way or another, and therefore are not concerned with nativism. Extended discussion on this topic could be found elsewhere (McClelland et al., 1986; Shepard, 1989).

3.4 Pattern recognition

Chapters above provide an overview of what connectionist networks are capable to achieve by mapping one set of patters onto another by encoding statistical regularities in connection weights modified by the network learning process. This chapter provides a discussion on the claim that networks are particularly appropriate for modelling behaviour, which entails that patterns are central to a variety of human faculties and connectionist networks are particularly appropriate for it.

Pattern recognition could be defined as mapping a specific pattern onto a more general pattern. Another type of mapping is *pattern completion*, which is mapping an incomplete pattern onto the same but completed pattern. *Pattern transformation* is mapping one pattern onto a different related pattern. *Pattern association* is mapping a pattern onto a different unrelated pattern.

In human behaviour, pattern recognition is most apparently evident in perceptions, where local classifications are combined into higher order patterns, which in turn serve as inputs for high-level recognition faculties and abstractions that are recognised by human languages. Thus, *categorization* does not only refer to semantically interpretable cognitive level of pattern recognition, but also to lower-level sensation and perception (McClelland, 1979).

3.4.1 Pattern recognition algorithms

The following sections demonstrate mapping abilities of connectionist networks.

3.4.1.1 Pattern recognition in two-layer networks

For a system to be considered capable of pattern recognition, it should display a consistent response to the instances of pattern presented to it. Two-layer networks are quite competent at such tasks and can learn to recognize the pattern using the learning rules discussed above. Moreover, while doing so the network would develop a good generalizing capacity. It not only will be able to respond well to input patterns seen in training, but also to patterns previously unseen – producing an output closely resembling the output of the similar known input pattern.

This however is somewhat different to the typical human learning process as the exposure is not restricted to the perfect examples but rather is an assortment of similar cases with varying levels of distortions. Therefore, the network was tested using the distorted inputs and output patterns. As a result, the network was able to provide a qualitatively correct response very close within the numeric values to a desired output – even to previously unseen patterns. Thus, the simple two-layer network is able of learning to recognize several input pattern categories, and can handle distortions in the pattern and respond to new patterns quite well and in a natural manner.

3.4.1.2 Pattern recognition in multi-layer networks

Mapping input patterns directly onto output patterns is not sufficient for some pattern recognition, and may require additional intermediate layers to facilitate the information extraction. One such early interaction model was designed to recognise visual patterns – four-letter words in a certain font (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). Input layer contained individual elements of letters, intermediate layer contained letters, and output layer contained four-letter words. The intermediate layer in this particular model is not a true hidden layer as the containing features were designated by the researcher rather than being extracted as the result of the network learning process. Instead, the intermediate layer was set up as a sort of an extra output layer, where the network is able to report on either the letter or the words depending on the task parameters. In a true multi-layer network with hidden layers there is usually little reason to report the hidden layer as it normally contains a sophisticated set of microfeatures extracted by the network that are not easily interpretable.

Even though the network was designed decades ago and before the back-propagation learning procedures were developed, it is able to accommodate in a very human manner a range of conditions such as low contrast and missing elements, and exhibit graceful degradation due to multiple soft constraints that the network is aimed to satisfy.

Another fascinating aspect of human pattern recognition addressed by the model is *word superiority effect* where letter recognition is improved when presented in a context of a word (Reicher, 1969). Explaining the underlying processing that considers the actual word in the course of letter recognition was something that presented a challenge for researchers. McClelland and Rumelhart's model (1981; 1982) provides one auspicious explanation where word recognition may affect component letter recognition. As described earlier, the network contains a layer of letter features, layer of letters, and layer of four-letter words. Each of the feature units is positively connected to those letter units that contain the features and negatively to those that do not. In the same manner, letter units are positively connected to those word units that contain the letters in the appropriate position, and negatively to those that do not; and word units are positively connected to the letter units that the words contain. Additionally, all competing combinations are negatively interconnected. When the input to the network is supplied by activating all the feature units for all four letters of the word, they excite the letter units to which they are positively connected, which in turn excite appropriate word units. Word units will send the activation back to the letter units, and the activation propagation in the interactive network will continue for a number of processing cycles. This activation propagation direction forward and in reverse is critical in our discussion of *word superiority effect*, as it shows how the letter layer is activated from feature layer and from the word layer in reverse direction. Thus, if feature units do not readily identify a word through the letter units, the word unit with the best fit for the features would activate those letters that are able to form a word. This allows the network to deal with inconsistent and partial data such as misspelled or unclear words. If there is one best fitting solution, the network will identify that word by activating the missing letter unit through many cycles.

If however there are a number of possible solutions that fit, network behaviour is even more interesting. The partial feature input would initially equally activate all letters that fit, and once the activations would start going in reverse order the word units would be able to exert activation onto letter units. Since certain words are more frequently encountered, they would carry higher resting activation and therefore would be able to inhibit other word units and therefore the appropriate letter units, eventually being identified by the network as a solution. Thus, this illustrates how a higher-level knowledge is able to exert influence on the lower-level letter and feature units, suggesting that additional layer that would consider word units in a context could have an effect on resting activations and therefore on the overall network behaviour. This demonstrates the ability of network models to *complete patterns* by predicting what is missing in addition to already discussed *pattern recognition* ability. A more advanced network design could use sensory inputs as lower-level units and theoretical or philosophical statements as higher-level units, which could with the help of learning procedures exert influence on lower-lever sensory units to facilitate the pattern recognition process.

The network performance illustrated here is remarkable, and shows how the task can be accomplished with the use of connection weights and activation functions rather than the use of rules. Since, more sophisticated multi-layer networks have been developed that make use of learning procedures such as back-propagation and proper hidden layers able to extract microfeatures from the input patterns useful in advanced information analyses.

Networks have also been used for semantic categories recognition modelling, as discussed next.

3.4.1.3 Generalization and similarity

One of the key performance faculties that the networks exhibit is their ability to generalize. Once the network is trained to classify the input patterns into certain classes, when presented with a previously unseen pattern, it will provide a response comparable to the response to a similar known pattern. What constitutes this very similarity is a fascinating contemplation. One explanation would rely on the properties shared between the two or more entities, which bring upon the philosophical argument that any two entities share an infinite number of properties. Thus, assessing similarity in terms of number of shared properties is deficient unless constraints are imposed.

Humans however are quite adept at judging similarity. Networks also have a very clear way of doing so – the similarity structure is an inherent component of the weight matrix. Similarity however is a complex concept, and does not necessarily have an objective measure or device capable of assessing it: if a network generalizes in a manner dissimilar to human reasoning, it is natural to assume a failure. However, it is crucial to consider the possibility that network may be capable to generalize on a different level, perhaps not readily comprehensible or qualitatively different. In fact, considering that current connectionist network architectures are rather quite simplistic compared to the neurophysiologic complexity of human brain, it should not be surprising to observe dissimilar behaviours in the networks and in the mind. Environment has been shown to play an important role in the developmental process of human cognition, and the connectionist network devoid of such experiences may be unable to comprehend entirely the sense of similarity in the same manner as humans do.

The fact that networks generalize following the same mechanism as the process of pattern recognition should be considered a benefit as it does not need to involve the extensive philosophical discussion briefly touched upon here.

3.4.1.4 Pattern recognition beyond perception

In connectionist networks, pattern recognition plays an imperative role at all levels of analysis – sensational level to reasoning – without clearly defined boundaries between the concepts of perception and cognition. In contrast, some symbolic theories (for example Fodor, 1975) consider symbolic processing as isolated from the sensational processing and is not regarded in terms of pattern recognition. Some other symbolists (for example J. R. Anderson, 1983a) are similar to connectionist networks as the pattern recognition does occur at all levels of analysis. Any system substantially reliant on pattern recognition is a prospective candidate to offer a plausible account for the *intentionality* of mental states. This notion is discussed in detail in the following section.

3.4.2 Intentionality of pattern recognition

In philosophy, intentionality refers to the notion that mental states have meaning and content. Intentional states are concerned with the phenomena that are outside the cognition, and it has been one of the more testing tasks in cognitive philosophy to describe how mental states become associated with the specific phenomena and acquire the intentionality. The difficulty revolves around the relation between the mental state and the external phenomenon that is unlike any other. If one person believes something about the other, the first person seems to have a relation to the other, and both need to exist for a relation to be true. However, the other person very well may not exist at all, and yet the belief could still be possible. For that reason, such a connection cannot be handled by the means of relation alone.

Trying to solve the intentionality issue with the traditional symbolic approach to cognitive modelling is particularly difficult, as the representations employed by the symbolists are formal and the best that can be achieved is describing the objects to which the symbols refer. This, however, only repositions the issue, as it is now necessary to explain how the symbols used in the description relate to the external phenomenon. The difficulty in explaining intentionality lies in finding a way to relate representational states to the real phenomena. One such approach is to consider the causal mechanisms that generate symbols in terms of the transmitted information and how the symbol relates the information about the object. Thus, when the symbol is activated without being caused by the referent, it is still by the object that would cause activation in normal circumstances. This framework however does not provide a plausible account of representing nonexistent objects. Another problem with the symbolic approach in trying to relate representational states to explain intentionality is that the arbitrary treatment of symbols, resulting in the inability to relate the symbol to the referent and the symbols becoming context-free. This is also evident from psycholinguistics where meaning of certain words is dependent on the context – much like referent may vary with the context. Explaining such causal relation with symbolic models poses a problem in trying to explain significant variations in intended referents depending on the use of the symbol within a different context. Employing a more complex system of symbols to account for varying contexts is one possible solution of addressing the issue. Another proposal (Barsalou, 1983) suggests that concepts are not fixed, but rather are construed each time from the individual elements appropriate to the context.

This proposal would be appealing to the connectionist perspective, where symbolic elements could be treated as microfeatures spread across the units in the network and additional input units could account for the context sensitivity. Connectionism proposes the occurrence of processing within the system is uninterrupted with the processes taking place in the external environment, avoiding the separation of the sensation and perception in symbol processing. Hence, the cognitive processing is position as occurring within the external environment, where individuals use skills and behaviours to interact with objects at varying level of abstractness. Pattern recognition networks capture the regularities in behaviours at different abstractness levels suggesting a good fit with the system and the external environment.

One of the key differences that separate connectionism from traditional symbolic approach is that the connections that represent the interface of the system with the external environment are not arbitrary, but rather are the result of the learning process where only the relevant connections are defined as a pattern that represents the interaction of the system with the external environment. Consequently, a two-layer pattern recognition network would modify the connection weights to reflect the input-output relation dictated by the environment directly, whereas multi-layer network in addition would encode the higher-order information such as microfeatures into the hidden layers. The input in most cognitive modelling networks however is specified by the researcher and does not incorporate the environmental parameters, and therefore would not provide a sufficient evidence for the claim that network representations are directly linked with the object. It is quite a common practise in other disciplines though (for example engineering) where the networks are supplied with the ability to receive a limited input about the outside environment, and in that case the representation is very much about the external object.

A point of the essence to be made here is that representations in the hidden layer are the result of network accommodating to the environment, and do not constitute causal connection with any sensory input which makes them arbitrary from the standpoint of system functionality. Connectionist learning systems are designed to perform specified tasks, which involves functioning in a certain external environment. The learning procedure defines a goal for the network (error minimization in determining output pattern to a given input pattern), and representations constructed in the hidden layer serve these goals by embodying the external to the system information for the system, thus making these representation about the external environment.

System response to a particular input is not necessarily context-free. Information used in unit activations may correspond only to a general body of information about the environment (contextual and other) in which the system is present; and, depending upon the goals set, system learns to identify it through the responses to the patterns. Therefore, system response is a complex combination influenced by numerous factors some of which are only partly related to task at hand and yet exerting influence on the general patterns of activation from within the system. This versatility allows the system to adjust the response in relation to other information available.

Connectionist approach to modelling cognition is able to provide knowledge about the intentionality of mental states, where representational values constitute the network's response to the input pattern. Since the network is adapted to the input pattern, the network state could be directly link to the external environment if properly connected (sensory input units). Sensitivity of the representations to the external and internal context makes even better of a case in attempting the explanation of the nature of representations.

Discussed earlier instance of explaining the mental states that represent nonexistent object could be resolved by the symbolic models with a relative ease (employing a symbol for nonexistent object), still unable to provide an explanation as to why the arbitrary symbol is linked in such manner. In connectionist networks, it is possible that the output pattern will be the result of internal network activity, thus providing an output that does not necessarily correspond to any of the input patterns and therefore would represent a nonexistent object. Such outputs would still be based on the featural elements defined by the system and for that reason be a representation of such objects and not the others.

This shows the important role that pattern recognition can play in intentionality. Related to this philosophical discussion is a string of research in cognitive psychology on the formation of semantic categories, which is discussed in the following section.

3.4.3 Categorisation with connectionist models

This section outlines the progress in the body of psychological research on concepts and categorization, and how symbolic and connectionist models could be of relevance.

In categorization, methods of symbolic and connectionist modelling does not differ to the high degree. It is possible to assign a symbol to the category, but even do the symbols could serve little purpose without some sort of distributed representation that relate the features to the categorical assignment – generally referred to as *exemplars*. For that reason, even within the symbolic approach to cognitive modelling categorization has been handled in a conceptually similar way to pattern recognition: exemplars assigned to semantic category according to their featural compound. Thus, a lot of research on categorization conducted within the symbolic approach could be transferable to connectionist networks modelling techniques. Both approaches acknowledge the primary dissimilarity between the pattern recognition and assignment of exemplars to semantic categories to be for the most part concerned with the featural level of abstraction: low level (strokes in handwritten word recognition task) versus intermediate level (possession of gills in animal classification task). Contrary to the symbolic models, distributed representation across features may be sufficient to represent the semantic category in a connectionist network, i.e. it does not require a designated symbol or unit to denote the category.

To demonstrate how the connectionist pattern recognition model is able to accommodate the classification mechanism, consider a two-layer network. Input units represent specific features of the exemplars, and the particular pattern across the inputs is a distributed encoding of the exemplar across the appropriate features. In the same way, the output units represent the distributed encoding of the probable categories, where the connection weights appropriate exemplars to the suitable category – uncharacteristic to the category features would have low connection weights with the category. Once presented with the exemplar, network would propagate the input activations along the connections, and each output unit would receive a summative activation from its inputs. Additive combination of features is a modelling technique not a uniquely distinctive characteristic of connectionist networks but rather is quite common to a whole class of characterization models.

The method of distributing categorical representations across features and their utilization have been a major research interest within the field of psychological categorization modelling. The classical view follows the philosophical analysis, where it is assumed that categories identify the sets that are defined by the necessary and sufficient conditions; and knowing those conditions constitutes knowing the categories. Consequently, the view suggests that all categories are processed in a comparatively similar manner and exemplars are treated equally.

In the 1970s however, fundamental changes were proposed by Rosch and others (for example Rosch & Lloyd, 1978) that challenge both consequential views. It was demonstrated that in class-inclusion hierarchy one level among others is the *basic level* and therefore is processed and acquired more easily; and categories have a ranking structure where some exemplars are recognized as better representatives of the category (Rosch & Lloyd, 1978; Rosch & Mervis, 1975). In addition, it was demonstrated that prototypicality played an important role across many information-processing faculties, and recognition and categorization of typical exemplars shows better results as far as time and accuracy.

The typicality of the exemplars to the category is judged based on the features shared. The features though do not necessarily follow the classical definition of the category, nor need they be common among all member of the category or be distinctive. Moreover, the typicality could be demonstrated among such obviously defined categories as odd and even numbers (Armstrong, Gleitman, & Gleitman, 1983), which indicated that typicality must depend on elements other than classical definition. Therefore, categorization should consider category definitions along with the typicality effects, which may still be insufficient to represent knowledge structure on categories in an adequate manner.

Early cognitive models adopted the classical approach to represent knowledge of categories, relying on logical statements and necessary and sufficient conditions. With similarly aims, semantic networks (for example J. R. Anderson, 1974; Norman, Rumelhart, & Group, 1975) implement highly localist structures where units encode semantic concepts interconnected by a small number of connections conveying the relations between those concepts. Hence, cognitive propositional model and semantic networks represent two approaches for symbolic knowledge architecture. Following the discussion on the prototype, the semantic networks approach could be adapted to accommodate the idea that mental representation revolves around the prototype rather than propositional logics that determine the category: some representatives of the category may have distinctively differing qualities than the others and therefore would lack the corresponding connections. To account for effects of typicality, semantic networks adopted the process of *spreading activation* (J. R. Anderson, 1983b; J. R. Anderson & Pirolli, 1984), which signified a significant breakthrough in the advancement of network models in the later years (having particularly high resemblance with the localist connectionist networks and both can account for typicality by the means of summative weighted features).

Prototype and abstraction models specify a somewhat different theoretical position where belonging to a category in exemplars is assessed based on their similarity to prototype. Multidimensional scaling of lists of features could be employed to represent conceptual frameworks that comprise of features with varying parameters, suggesting a characteristic rather than a defining feature (including both continuous and discrete features in the model). *Exemplar models* that followed inherited the probabilistic view of categorical structure, but contrary to the prototype models, the category is represented in more detail by the exemplars rather than the prototype (Medin, 1989). Individual featural representations for every exemplar are stored and weighted, and therefore similarity computations may involve weighted summative feature analysis as in connectionist networks.

In categorization tasks, exemplar models present a direct competition to connectionist models, providing an alternative way of distributed representation across features in prototype extraction and categorization tasks. The distinctly dissimilar assumptions regarding storage and processing in exemplar and connectionist models may result in essentially different result of computation. Connectionist networks retain information about exemplars only if this has an effect on the weight matrix, and shift to prototype extraction for similar exemplars as the exemplar numbers increase (retaining as much information about exemplars as possible), using feature vectors for temporary activation patterns. Exemplar models store information about particular exemplars as feature vectors, which are used to compute the prototype. Ability of connectionist network to model the information about exemplars and prototypes within the same framework may offer certain advantages – especially considering the much broader application in modelling a variety of cognitive tasks, and not only categorization.

Categorization, as proposed by (Barsalou, 1983), could also be interpreted not as a stable mental grouping of fixed entities stored and retrieved from memory (Rosch & Mervis, 1975) as required but rather are produced as the particular task is being performed – idea supported by the fact that people are able to construct new categories upon request. The emergence of these concepts then relies on vast amounts of continuous knowledge stored in long-term memory, which is used to form temporary relevant to immediate context concepts in working memory. Unlike symbolic models, connectionist framework is able to interpret these findings in the following manner: concepts could represent stable patterns of activation that determine further processing. However, on a different occasion, the resulting patters could be altered due to activity elsewhere even using the same weights. This may represent the continuous knowledge in long-term memory.

Thus, pattern recognition capacity of connectionist networks is able to perform categorization tasks, exhibiting typicality and task-sensitive variability effects – some of the requirements that must be met by a successful model of human categorization.

3.4.4 Pattern recognition in mental processes

Human cognitive abilities go far beyond the relatively lower-level tasks of perception and semantic categorization and classification: phenomena could be contemplated *sans* the actual perception taking place, a person could perform a hypothetical planning of future behaviour, and other higher-lever tasks usually construed in terms of performing logical inferences on symbolic representations. On the condition that pattern recognition actually underlies much of cognitive faculties that necessitate reasoning, connectionist framework would be able to provide a plausible account of higher-level tasks in similar manner as in the case with the lower-level tasks already discussed above.

One possible structure that may enable the relation between the pattern recognition and all-inclusive account of cognitive ability is to utilize the stable state representing one pattern as an input for the next level pattern recognition system of higher order. Thus, the reasoning steps of cognitive performance could be represented as a sequence of multiple levels of pattern recognition. The work of (Margolis, 1987) supports the proposition that human function could be explained in terms of pattern recognition, where humans decompose complex situations by recognizing something and invoking the most suitable to the situation pattern, which is then used to recognize something else and therefore modified to reflect the situation to a greater degree, constituting the learning process. In addition to the process making a judgement through the process of reasoning, the process of reasoning why it is the case. This review of the process of reasoning and making a judgement is in itself a separate pattern recognition event that may provide and insight into the pattern recognition process on a broader spectrum, and result in a modification of the pattern recognition mechanisms based on the acquired learning.

The two arguments that (Margolis, 1987) offers to substantiate his claim revolve around the seeming human limitation in logical and statistical faculty, and the example describing the ability to adopt new scientific paradigm by some scientists and resistance to such change by other. The first, based on the work of Tversky and Kahneman (1973, 1974), describes the human difficulty to perform an accurate statistical probabilistic evaluation – the fallacy that occurs according to Margolis due to the inability to elicit appropriate pattern and is, therefore, not a statistical but rather a pattern recognition error.

The second example is drawn from the adoption of the evolution of scientific paradigm. Margolis argues that the ability to adopt and embrace the new scientific paradigm requires learning to recognise new patterns. Developing new patterns often requires the abandonment of the old established patterns and may result in temporal deterioration of performance. This however does not constitute that cognition comprises exclusively of pattern recognition faculties – merely suggesting connectionism as one plausible explanation of cognitive function, and without the use of symbolic rules. For a discussion on the mechanism that illustrates how higher-order cognitive tasks may be performed using the pattern recognition function rather than logical reasoning of a symbolic system please see Rumelhart et al. (1986).

3.5 Knowledge representation in connectionism

The propositional knowledge representation that revolves around the idea that knowledge is expressed and therefore can be transferred in propositions such as sentences is accepted in a quite intuitive way. It is generally accepted by the cognitive science disciplines (cognitive psychology, artificial intelligence, etc.) that knowledge is represented in propositions: it is transmitted by books and lectures that consist of sentences formed from mental sentence-like structures. Many report a kind of an internal dialogue when describing their thought process, and traditional information-processing models of cognition and language share the assumption that propositions are what is processed. Connectionism on the contrary, poses a challenge to this general assumption regarding the knowledge representation, and networks are able to encode the knowledge in a qualitatively different manner without the necessity to employ propositions, effectively rendering the concept of propositions for cognitive modelling unnecessary.

3.5.1 Knowledge representation in Cognitive Science

Earlier efforts to model cognitive representations have relied on unstructured declarative statements arranged according to predicate calculus rules: predicates followed by a number of arguments that were seen as basic conceptual parts. Propositions were connected through the rule of repetition and involved recurring concepts. Further research by psychologists and artificial intelligence researchers revealed the need for higher order structures sufficient to organize the propositions, initially called schemata (Rumelhart, 1975) or frames (Minsky, 1975). Schemata represent the structured framework of knowledge where propositions are allocated appropriate location. Once activated in the course of cognitive process, schemata offered a response by default unless contradicting information was available as a substitute and an update of the schemata tailored to the immediate environment. Schemata and other higher order structures of proposition organisation introduced to the propositions representation a limited ability and characteristic of a pattern recognition system, where some knowledge parts are semantically connected to other knowledge parts to make possible further knowledge processing.

In 1970s, one of the challenges to dominant at the time propositional approach came from the researchers that argue that knowledge was represented as images – analogue and iconic in form in contrast with the abstract and arbitrary propositions (J. R. Anderson, 1987). These distinctions between the literal pictures and non-literal representations in visual and spatial form prompted the emergence of multi-code models (J. R. Anderson, 1983a; Paivio, 2013) with modality-specific knowledge representation, where visual information is encoded as images and literal information as a verbal code – challenging the claim that propositional representations are sufficient to encode all knowledge. Further studies with clever design were able to establish linear relations between the analogue dimensions and the response reaction time (Kosslyn, 1980); and even suggested that analogue representations are useful in carrying out inferences by employing mentally ordering objects in spatial array (Huttenlocher, Higgins, &

Clark, 1971), challenging the purely propositional knowledge representation further.

Until the re-emergence of connectionism, cognitive scientists held models of pictorial knowledge representation as alternative to propositional models – as complementary only and useful in representing certain types of knowledge rather than able to completely substitute propositional representation models. Connectionism on the other hand, aims to provide a plausible account to some or possibly all cognitive performance without any use of propositional representation, thus occupying an opposite position to the traditionally established approaches.

While considering connectionism as a model of cognitive performance, it is useful to consider the concept of cognitive performance itself outside the propositional representation theoretical framework of knowledge representation.

3.5.2 Types of knowledge

The distinction between the *knowing how* and *knowing that* developed by Ryle (2009) is based on the human ability of not only knowing certain facts, but also on knowing how to perform certain activities. The capacity to possess both types of knowledge therefore is also different: *knowing that* requires the storage and consecutive retrieval of the specific or relative proposition from memory. Whereas *knowing how* may require a specific knowledge of the process and the control of associated perceptual and motor systems necessary to complete the activity successfully, such as planning, execution, monitoring, etc. Thus, the propositional knowledge represents only a portion of human intelligence and is not primary; and it is often the success rate in performing certain activities, both physical and cognitive in nature, we are interested in while assessing the level of intelligence. In fact, the intelligent practise or and theorizing itself is but one of the activities that could be conducted in an intelligent or a stupid manner.

The behaviouristic account of Ryle's *knowing how* consists of the disposition to perform the activity in the appropriate circumstances, without specifying the internal mechanisms involved. In cognitivism, it may be appropriate to attempt an explanation of *knowing how* in terms of learning how to perform the activities and what mental activities obtaining such knowledge involves. Cognitive science provides an account of *knowing how* from the rule-based systems point of view, where it is referred to as *procedural knowledge*. People often learn how to perform new activities from others by receiving verbal instructions on what to do, thus receiving procedural knowledge that enables them to perform the said activity. As such, procedural knowledge is rule-based and propositional, specifying the set of actions to be taken – for instance generative grammar in linguistics (Chomsky, 1957), where the rules are used as abstract representations of the competence level rather than a performance model. Verbal instructions alone however are not sufficient to perform the behaviour in a satisfactory manner, and require practice – actual or mental (Newell, 1994). Even though adapting the cognitive rule-based models initially developed to work with the propositional knowledge to accommodate the procedural knowledge has shown considerable success, it does not necessarily provide an explanation of *knowing* *how* in qualitatively different terms from the knowledge based on declarative propositions.

Connectionist networks on the contrary, are not ordered in strings but rather consist of interconnected units, and in many instances these units cannot be straightforwardly interpreted in terms of symbols. The *knowing how* may refer to the propagation activation in the network models, and therefore is more like a dynamical processing in connectionism rather than a sequential application of propositional rules.

3.5.3 Expert knowledge

Expert systems received a considerable amount of interest from the cognitive psychology and artificial intelligence researchers over the years. Many different tasks have been studied extensively, involving such complex activities as playing chess, medical diagnosis and other (J. R. Anderson, 1981). The typical approach aims to formulate a set of rules capable of achieving performance levels comparable to a human expert through interviewing the experts and surveying their methods and processes, which is then encoded as a computer programme that simulates human expert performance. Many expert systems show high levels of competence both theoretical and applied, which supports the notion that it is possible to incorporate the *knowing how* into the propositional systems initially designed to provide an explanation for *knowing that*. Not everybody is convinced however, and Dreyfus, Dreyfus, and Athanasiou (2000) after carrying out an extensive analysis of human skill acquisition and performance concluded that expert systems approach is inadequate in simulating the human expert performance, as expert systems are inherently limited in the level of performance they are potentially able to achieve. Based on their analysis, a five-level scale of skill development is proposed, where only the highest levels manifest the true expertise:

1. Novice.
2. Advanced beginner.
3. Competent performer.
4. Proficient performer.
5. Expert.

Dreyfus, Dreyfus, and Athanasiou (2000) argue that the work on expert systems is capable of addressing only the first three levels with the symbolic modelling where the major cognitive tasks can be grouped into assessment of circumstances, choosing the appropriate response and managing the rules to accomplish the objectives. Developing additional cognitive tasks of the same level or combining them to produce ones that are more complex would not be sufficient to advance the competent performer to a level of an expert.

3.5.4 Vision knowledge

Knowing how to see is generally considered such a basic function that many philosophers did not believe it may actually require knowledge to execute. Sensation and perception was considered in terms of evaluating the *observation sentences* to assess the truth-values, and the truth determination process considered unproblematic as a simple visual recording mechanism. This assumption however was challenged by Kuhn claiming that what we see depends upon what we know: someone familiar with FMRI would see FMRI machine for what it is, whereas somebody unfamiliar would see it as ‘some sort of a device or an assembly’ instead. Thus, the epistemological threat to the objectivity of the scientific process is revealed if the theory determines what the researchers will see during the observations, effectively introducing relativism into science. What is applicable here though is the notion that perception is a learned function and therefore relies upon knowledge that is not represented propositionally. Hanson (1958) accumulated a body of knowledge against the viewpoint that perception is simply a recording mechanism, and argued that all people rather than all seeing the same thing, each individual sees it from one of the dimensions first and then may adopt to see it from another dimension. Thus, FMRI technician would see the FMRI machine for what it actually is first, whereas layperson would first see it as some sort of a device and can then make an inference; which is dependent upon the learning, as to layperson needs to learn what the FMRI technician knows first before being able to see what the technician sees.

What remains to be explained now is the process that determines a mechanism for learning a set of propositions to facilitate perception, i.e. for perception system to see what it would not be able to otherwise see as a result of learning. This of course cannot be the bottom-up process of inference. Kuhn (2012) also supported the notion that perception needs to be tuned to the discipline and in case of significant shifts as a course of scientific progress, it is necessary for scientists' perception to be re-learned. This suggests that at least part of practicing science consists of *knowing how* to perceive objects and events. Thus, the actual process that describes in what manner does *knowing how* have an effect on the perceptual system, and how this type of knowledge is acquired, remains unsolved. Connectionism may be able to offer a way to resolve this. One of the features of connectionist pattern recognition system is the ability to perform even in situations with altered variations, as precise matches are not necessary between the current and already learned patterns. Non-obvious subtle regularities (often not easily describable in words) are extracted to identify exemplars and prototypes and are stored in connection weights – a nonpropositional way to encode the necessary for the task knowledge.

3.5.5 Logical inferences

Symbolic perspective designates logical inferences to be a lower level cognitive ability. The normal process of learning the logical inference rules for symbol manipulation consists of learning the rules and learning how to apply them as part of subsequent practice. Practise improves the outcomes, but the process does not become flawless and mistakes still happen. One possible explanation suggests that some rules may be learned incorrectly as separate rules are collapsed into one general rule later to be split into distinct separate rules. As a result, the general rule may be still incorrectly utilized on some occasions. This can be modelled by attaching the probabilistic parameters to the utilization of the rules, where learning is expressed in terms of changing the said parameter (J. R. Anderson, 1983a). Observations reveal however, that learning how to apply the rules is an essential part of the overall learning process as it develops the *knowing how* part in addition to *knowing that* – the explicit rules. Rule-based models are usually designed to incorporate certain functionality of pattern recognition. Would it perhaps be possible to eliminate the rules entirely, leaving the connectionist network and pattern recognition to do all the work?

One of the challenges in developing a simulation model of consumer behaviour lies in identifying the features and information that is used in the decision-making process – one option is to rely on the established and developed theory that specifies what is necessary.

Important to note that networks may require large number of training trials due the fact that network models normally start from zero (tabula rasa) – unlike humans that possess certain prior knowledge and training that may be applicable to certain extent. Moreover, networks, alike to humans, are capable of attaining high performance values as a result of repeated learning and error corrections in the course of pattern recognition activity not reliant on proposition-like rules. Thus, a distinctive differentiation between the *knowing how* and *knowing that* is made apparent in the ability to apply pattern recognition processes to linguistic symbols.

3.6 Summary

In this chapter, the arguments against the connectionist networks were discussed that outline the inadequacies of the connectionist networks to model cognition as opposed to the symbolic models. In response, the three connectionist approaches were presented that include: (1) the approximationist approach that argues that network models do provide a more accurate account of cognition than symbolic models, (2) the compatibilist approach aimed at building symbolic architecture into the networks, and (3) the external symbols approach. In this paper we are for the most part concerned with the kind of combination of the first and the third, as they offer a novel plausible explanation of cognition opposing the traditional symbolic approach.

4. Methods

This section provides an overview of the research questions and research methods employed. The data and analysis are described. The modelling approach and variables used are explained and justified, and research process is outlined in a sequential manner.

4.1 Overview

It is important to establish the boundaries of this research project, and discuss the overall goals it is set to achieve. Firstly, as further discussed in the following sections, the research objectives are set to explore the field of consumer behaviour, a primarily positivistic field of study (P. F. Anderson, 1986b), and attempts to model and examine the underlying architecture of the consumer decision-making process employing connectionist NNs models. It is argued here that utilitarian and information reinforcement are latent emergent variables that are represented by the input items and thus need to occur at the level higher than input level of independent variables (Foxall, 2009). Traditional methods such as regressions do not have any levels other than *Input* and *Output*, whereas connectionist network structures are able to incorporate a number of levels as *hidden layers* – where the utilitarian and informational reinforcement should exist conceptually as emergent concepts and representations. This is then principally an attempt to develop explanatory modelling that would allow examination of such higher-order attributes, and potentially offer a method to evaluate and approximate utilitarian and informational reinforcement quantitatively.

It is important not to overlook the predictive capacity of the model, and its ability to extract important patterns from the data, together with other dimensions considered as well, such as explanatory dimensions, including both descriptive and prescriptive application (Bryman & Bell, 2007). The overall context for the extended discussion here is of course the extension of theoretical framework of BPM to incorporate connectionist view as one possible direction to go forward.

To provide a comprehensive account of this research process, this section will focus on the following: (1) research questions and hypotheses, (2) philosophical position, (3) theoretical justification and evaluation of the approach and possible alternatives, (4) research methods, which includes a sequential account of research process, and (5) a concluding remark.

4.1.1 NNs models and linear models

Linear regression is undoubtedly one of the most widely used and important tools to describe possible relationships between variables in behavioural science. Many factors account for such widespread adoption – the seeming ease and intuitiveness of interpretations certainly being one of them. Linear regression models are usually fitted using the least squares approach designed to minimize the lack of fit, allowing either to quantify the relationship strength between variables or to develop a predictive model as a result. One of the most common applications is trend line estimation in time series data to show change over time

– simple technique that does not require a control group or sophisticated experimental design. Furthermore, predictor variables are often intuitively transformed to improve the function fit, making linear regression an exceptionally powerful inference method indeed – as is the case with polynomial regression that can be too powerful and may often show tendency to overfit the data. Employing the interactive variables is able to further improve upon the modelling results, providing the possibility to examine nonlinear relationships. Even so, considerable difficulties may be encountered when interactive variables are employed to examine the relationships in large datasets with many variables, and when relationships between three or more variables are examined, as the number of interactions increases exponentially, and readily becomes impractically large. Given *n* predictors, the number of items in a linear model that includes a constant, predictors, and every possible interaction is 2^n – with only 10 variables for example, total number of only 2-variable (excluding those of 3 and more) interactions to examine and evaluate is 1024, and the selection process is very tedious and manual, and oftentimes impossible. As a result, researchers may examine only a few interactive variables that first come to mind, or none at all. Another issue is the possibility of running out of degrees of freedom.

NNs, on the other hand, are inherently designed to examine dynamically all possible interactions within the data during the learning process. All interactions that carry predictive capacity are captured in the final network architecture by a learning algorithm, and pruning methods systematically simplify the network to expose the core explanatory architecture by removing the connections that do not offer sufficient predictive capacity.

4.1.2 Architecture of NNs models

In the simplest form, where the number of hidden layers is set to zero (that is a NNs model where the input layer is connected to the output layer with no hidden layers between the two); NNs model develops a structure similar to the structure of a logistic regression. As a result, the coefficient numerical values of logistic regression would be identical to those of the weights in NNs model. Thus, referring to the common 'black box' argument, it should be clear that NNs models in the very least are able to provide level of explanatory capacity equivalent to those of traditionally employed linear methods such as logistic regression. This however has already been explored in greater details elsewhere (for comparative analysis please see Greene, 2011), and in this paper NNs models of higher structural complexity are examined.

In the high-level task of pattern recognition while examining complex behaviour phenomena, linear models could only be useful in explaining linear relations. For the purposes of the present discussion however, this would be insufficient as consumer behaviour and the process of decision-making in a modern market and socio-economic environment is without a doubt a very intricate and multifarious phenomenon composed of a multitude of interrelated developments, where simple changes in one part of the system are able to produce complex effects throughout. It has been indeed a common practice to attempt to decompose the larger phenomena and isolate the process into individual elements for the following analysis controlling for all other variables. The learning thus obtained could then be propagated to the higher level of the process. This method however is very inefficient and poses a serious scalability problem – that is of course in addition to the limitation concerning the ability of researcher to identify the individual parts of the process correctly (the task some believe to be impossible). A better method would be to examine the relations between all components simultaneously.

NNs are able to examine all variables and account for nonlinear relations within the data once the hidden layers are introduced into the model structure. This results in high predictive ability, but also the weights could be examined for explanatory purposes and are able to provide an insight into the intrinsic nature of the process. Consumer decision making is an intricate continuous behaviour exhibited by persons that NNs seem to be particularly suited for as a method of analysis for a number of reasons. First, a NNs model framework as a method of analysis resembles physiological inner workings and structure of a human brain – making it a particularly good fit to study human processes. Second, connectionism (the theoretical framework of NNs) is a set of approaches in the fields of artificial intelligence and cognitive psychology that is particularly suited for modelling behaviour as the emergent processes of interconnected networks of simple units from the conceptual point. The hidden layers and nodes that are developed in the process of training a NNs model (NNs models are repeatedly fed data and adjust the weights in the process up to a point of equilibrium where the model cannot improve anymore – method commonly referred to as training as it indeed resembles the process of training in the traditional sense) are not like input and output variables that come from the data, but could rather represent underlying abstract concepts identified in the process of training that play a major role in explaining the relation between the input and the output layers (independent and dependent variables).

This paper will contemplate the idea of interpreting the NNs models number of hidden layers and nodes and weight values in attempt to provide an explanatory account of consumer behaviour. Previous findings will be summarized and synthesized, and original models developed and assessed (both predictive and explanatory capacity).

4.2 Research questions and hypotheses

The discipline of consumer behaviour encompasses contributions from a number of complementary fields of study, including psychology, philosophy, marketing, and economics (Bashford, 2009; Calder & Tybout, 1987; Holbrook, 1987; McKee, 1984; Pachauri, 2002). It is a common practice to produce research which is highly quantitative in nature (for example Cornwell et al., 2005; Cunningham, Young, Moonkyu, & Ulaga, 2006; Güneren & Öztüren, 2008; Lu Hsu & Han-Peng, 2008; van Kenhove, Vermeir, & Verniers, 2001; Watson & Wright, 2000), and is also the case for this project.

As discussed above, the central aim of this research project is concerned with extending the theoretical framework of BPM into the realm of Connectionism with the help of NNs models, assessing the ability of connectionist models to predict and explain the underlying psychological factors that influence and drive observable consumer behaviour. One way to operationalise this is to assess the capacity of a connectionist model to predict and explain the consumer disposition to pay more or less for a unit of product they eventually receive.

Therefore, the hypotheses are proposed as follows:

> *H1: Artificial neural network models with pruning offer means to simplify network architecture, while maintaining a level of predictive capacity comparable to unpruned neural network models*
>
> *H2: Consumer behaviour models based on connectionist framework offer means to examine the latent or emergent variables that represent complex consumer behaviour structures, which traditional linear models such as logistic regression are unable to elucidate*

The ability of connectionist models to develop the latent variables employing the distributed representations during the learning phases is given a particular attention in the discussion chapter, as this capacity offers unparalleled opportunity to develop this project further.

4.3 Philosophical position

There are a number of ways to obtain what we generally recognise as knowledge of consumer behaviour – ranging from a simple observation to a controlled laboratory experimental work (1988). As such, researchers tend to consider certain methods more suitable than others as applied to study particular phenomena. It is therefore important for the purposes of present research project to disclose and deliberate at least some underlying philosophical assumptions and perspectives generally adopted here. For illustrative and comparative purposes, the key philosophical aspects of each of the perspectives are juxtaposed and the manner in which they may influence the general direction of research are discussed: underlying ontology, epistemology, and axiology. Potential challenges that either of perspective presumes are identified. The section is then concluded with a summary remark.

4.3.1 Consumer behaviour position

For the most part, it could be considered a general knowledge that the field of consumer behaviour is predominantly lies within the domain of positivism (Marsden & Littler, 1996; Prus & Frisby, 1987).

As early as 1690 Locke (reprinted in 1997) argued that the method in social sciences should follow the same principle as it is the case in physical sciences, allowing for variances in prediction accuracy of course due to the obvious complexities of the subject matter. It was hundreds of years later the concept was adopted and further developed as School of Positivism by the likes of Quételet, Saint-Simon, and most notably Comte.

As proposed by *The law of three stages*, phenomena are to be explained through religion, metaphysics, and positivism, employing scientific laws, reason, and empirical data (Bernard, 1995). In modern positivism, Comte's original ideas are still present in the following form: scientific method is the optimal approach to generate effective knowledge with a reasonable degree of control, which can be used to improve the general human existence. *Logical positivism* is one of the later developments by Vienna Circle, also recognised as *logical empiricism* and *instrumental positivism*, stipulates that social science is to become a purely statistical exercise (Fullerton, 1987; Hunt, 1991).

Durkheim rejected most of Comte save the method, which was retained and advanced to establish methods and techniques for scientific research (Durkheim, 1964). Later alternative perspectives emerged as a result of critiques of positivism by Popper (1959), Kuhn (1996), and Foucault (1995), one of which is *relativism*.

In the following paragraphs, the two perspectives are compared and discussed in further detail.

4.3.2 Ontology

Ontological assumptions revolve around the concepts of what constitutes the nature of reality and of social beings.

4.3.2.1 Positivism

In positivism it is generally assumed for a single objective physical reality to exist, independently of one's perception (Peter, 1992).

Again, irrespective of individual perception, a single objective social reality is said to exist. Reality is composed of parts and interconnections, and it is separable and detachable; and it is possible to measure reality in a valid and reliable manner. Thus, to achieve a greater understanding of the subject examined, it is possible to control some of the other variables; and while individual inquiry may only be able to provide an estimation of reality, collective effort should allow developing a greater understanding and representation of reality. This implies the concept of decomposition of complex phenomena, where parts of the phenomena are taken out of the complexity and examined individually in isolation one part at a time to determine the intricate relationships and broader context.

A number of assumptions can be identified in positivism when it comes to the nature of social beings. It is possible to interpret human behaviour as reactive: in behaviour analysis for example behaviours are said to be reinforced by external factors acting upon the individual, and reinforcers systematically change the frequency of reinforced behaviour (Hildum & Brown, 1956; Insko, 1965). Similar predeterminism by outside factors is suggested by a cognitive view: rather than reinforcers directly affect behaviour it is the internal rationalization of the person that allows to make an optimum decision based on the available information and experiences previously processed by the individuals (Slovic, Fischhoff, & Lichtenstein, 1977).

4.3.2.2 Relativism

In relativism, subjective or objective reality exists only relative to the *relativiser*: a person, a theory, or other. If the relativiser is a person, it may be a perceptional judgment that is *relativised* and claimed to be *true* for that particular person – referred to as *semantic relativism*. A perception which is claimed to be *real* for that particular person is referred to as *ontological relativism*, and these two types of relativism are not always explicitly distinguishable (Long, 1998). When all judgments a person makes are true for that particular person, it can be described as *full semantic relativism* – thus not only the truths but also the whole reality is relative, which is in direct contradiction with positivism. By the same logic, what leads a person to construe a judgement as truth also lies within this same reality (Hunt, 1990). Therefore semantic relativism entails a version of ontological relativism, where that which constitutes truth for the person is within the full semantic reality (Nola, 1988).

Perceptional experiences could reveal ontological relativism as well: every person has their own perceptual experiences that no other person has had, and these individual perceptual experiences are subject to a prompt change. This would suggest perception dependency on the perceiver, or that every person exists within his or her own world of perceptual experiences. In the context of scientific inquiry it could then be claimed that concepts exist only relative to certain scientific theories, paradigms, and scientific frameworks (Nola, 1988).

Constructivism could be said to be interrelated with relativism in the context of what accounts for existence of phenomena: if researchers play an active role in creation and unravelling of scientific theory, their activities could be interpreted as relativistic in terms of what objects are considered by the theory (Goulding, 1999). It is then possible for positivists to agree that researchers indeed play an active role in the process – at the same time rejecting the implication that researchers actively impact the objects of theories (Nola, 1988). Socially constructed realities could then be seen as a product of relativism, as they are dependent on other entities – contrary to the positivist view of a single, objective reality.

4.3.3 Epistemology

As any other social science, positivism and relativism hold certain assumptions around the concepts of what constitutes knowledge, the nature of causality, and position of the researcher relative to the subject of inquiry.

4.3.3.1 Positivism

Positivism ultimately aims to derive the abstract generalizable laws that could be applied to a wide range of individuals, situations, and phenomena. In other words, positivists focus on determining the generalisations irrelevant of time and space that are context-free as much as reasonably possible. Single events do not hold any particular value unless they can be extended across systematic or sequential generalizable instances (Bernard, 1995).

The meaning of causality plays a central role in positivism. Fundamental to the underlying goals and values of the perspective, it is generally assumed that clear linkages could be established between behaviour and prior events that led to it. The deterministic nature of social beings is closely interrelated with the concept of causality, as external events that act upon the individual are presumed to affect individual behaviours or serve as affective factors otherwise (Bernard, 1995).

It is generally recognised in positivism that it is possible for researchers to largely remain outside the research, and consistently strive to distance oneself from the subject matter not extend any significant influence over the experimentation criteria. The researcher, drawing upon expertise and research methodology, is capable of manufacturing a hypothetical observation deck and, hidden by impartiality and detachment, remain objective. The position of the researcher in relation to the subject of study is then naturally assumed to be largely detached (Hunt, 1993).

4.3.3.2 Relativism

A number of ways illustrate epistemological relativism. Firstly, as suggested above, epistemology is inherently relativistic at least to some extent in a way that *what is known* is relative to the underlying theory, framework, culture, person, and so on. Therefore, *what is known* in relativist terms is dissimilar to what constitutes knowledge in positivist terms, as relativist knowledge is relative to something or somewhat rather than being seen as an absolute concept in positivism (Nola, 1988). It is then should be possible to reinterpret a positivist statement of *What is believed to be true, could be true or false* using the relativist terms as follows: *That which is believed to be true is true for whoever believes it to be true*.

Secondly, epistemological relativism occurs naturally due to inherent variability in perceptual capacities, and the concept of incommensurability develops at the level of observation in relativism – as opposed to positivist view that stipulates the possibility of objective observation (T. S. Kuhn, 1996). Feyerabend's (1975) view on methodological anarchism draws upon precisely this notion of epistemological relativism where it is applied on the level of methodological procedure. Epistemological relativism assumes that any observation follows a presupposed theoretical framework and therefore is inherently relative to that particular theory, and thus unable to generate objective data which would enable extrapolation of universal laws or rules (Nola, 1988).

Thirdly, one of the largely popularised Fayerabend's (1975, 1978, 1987) arguments claims there are no methodological rules. He challenges a single perspective approach on the grounds that it would effectively limit the scientific inquiry serving as a framework of constraints, whereas theoretical anarchism to the contrary would facilitate the scientific progress. Adhering to the rules of a single given perspective not only does not aid, but at times may even hinder the scientific process. General scientific description cannot be described through philosophical consideration, which would make it impossible to devise a method to differentiate between science and pseudo-science (Nola, 1988).

Feyerabend (1975, 1978, 1987) goes as far as to say that a universal scientific method does not exist, and any form of inquiry does not require a predetermined methodological process. Scientific reason does not need to follow any prescribed form of regulation that specifies a privileged perspective, as procedures designed to establish a prescriptive system would result in offering different incomparable ranking systems none preferable to the other. Admittedly, it may be possible to achieve this within a certain constraint – no universally applicable method could exist however. This key concept lies in direct contradiction with positivist notion of universal laws that could be applied to general phenomena.

4.3.4 Axiology

What constitutes value by either of the perspectives is discussed in the following paragraphs.

4.3.4.1 Positivism

The overall goal of positivism is prediction through derivation of universal laws that may be able to explain behaviour (P. F. Anderson, 1986a). The understanding of phenomena is tied to systematic demonstration of underlying associations between the variables selected to represent the phenomena. The accurate identification of these variables and antecedents that are related to the dependent variables is central to positivist perspective, as it would offer a degree of certain predictive capacity to be developed based on the results of the analyses (Kerlinger, 1964).

4.3.4.2 Relativism

In relativism, the methodological criteria are selectively employed by the scientific community and interpreted in response to empirical and social factors – as opposed to the criteria in hypothetico-deductive methods generally employed in positivism (Holbrook, 1989; Holbrook & Hirschman, 1982; Holbrook & O'Shaughnessy, 1988). Relativism is quintessentially descriptive, as relativists strive to develop a full comprehensive account of the phenomena rather than extrapolating universal law-like relationships that could be applied to general phenomena. It is not associated with any one particular method of inquiry – the theoretical framework is based upon empirical and qualitative evidence, and may include data of qualitative, quantitative, historical, and social nature along with any other sources that could prove useful in the attempt to develop a comprehensive representative account of the phenomena (P. F. Anderson, 1983, 1986a, 1988a, 1988b; Lutz, 1989; Siegel, 1988).

4.4 Research methods justification

In this section, the method of inquiry is reviewed and justified against alternatives.

4.4.1 Assessment of the quantitative method

Some of the limitations of quantitative method that researchers should consider may include inappropriate application of statistical methods and techniques to carry out the analysis, which may in extreme cases reduce the research project to the level of a purely statistical exercise that does not carry any other purpose. Highly technical and demanding, quantitative methods could be exhausting not only in terms of computational resources, but may also be limiting for the researcher in terms of the skills. This could lead to the quantitative method being susceptible to mistakes and inaccuracies that may result in errors and drawing of wrong conclusions altogether. On a separate note, there are some researchers who are not comfortable with research findings that derive meaning from numbers, employ quantitative and standardized data, and statistical modelling (Saunders, Lewis, & Thornhill, 2009).

Provided the researcher's philosophical position is aligned with the quantitative method, many potential limitations outlined above could undoubtedly be interpreted as an advantage. The concepts that deal with validity, reliability, and generalizability are often better accounted for by quantitative method as they are inherently imparted in the design (Ghauri & Grønhaug, 2005). One other major significant advantage is that any academic research publications are expected to be described following the positivist method, irrespective of the actual perspective employed (Wolcott, 2002).

4.4.2 Theoretical justification

Considering the nature of research questions set in this research project, it could be argued that no other than positivist theoretical framework may be suitable: it is unlikely any other researcher than positivist would even consider examining the capacity of connectionist framework to accurately predict consumer behaviour for example. Moreover, it could be said that research questions discussed here are inherently interrelated with the core values of the positivist theoretical position and thus form and inseparable part of it, as other than positivist researchers would not concern themselves with such positivistic notions as predictability, and would not therefore choose predictability as a central measure of research questions to begin with (Alvesson & Deetz, 2000).

It is uncommon however to encounter consumer behaviour research that is not based on quantitative method (for example see Haigh & Crowther, 2005; Holbrook, 1989; Kaynak & Kara, 2001; Kehret-Ward, 1988; Kumcu, 1987; O'Shaughnessy, 1985; Sanders, 1987). In fact, researchers continuously engage in an ongoing debate on whether positivism and quantitative method are appropriate at all to study consumer behaviour. Haas (1987) for instance argues that the subjective and social nature of consumer behaviour are obscured by positivism and its objectivity. He would argue that human consumer behaviour is above all a social process, and meaning derived from the interaction of individuals should be preferable to that which is based on a quantitative method: individuals do not respond to stimuli in a mechanistic manner as prescribed according to the theoretical assumptions of the school of behaviour analysis, but rather construct activities in a meaningful intentional manner. Behaviour is individualistic, and is a product of perceptions, interpretations, and judgement statements within a certain context, and therefore requires to be explained through the perspective of the individual (Haas, 1987). Denzin and Lincoln (2000) go as far as to state that all research is essentially interpretative, as it is fundamentally conducted in a manner that reflects the researcher's weltanschauung – a set of beliefs and feelings about the world, and how to understand and study it. Therefore, an interpretative approach would be preferred to positivism, being capable to produce a profound and thorough understanding of behaviour.

4.4.3 Alternative approaches

Considering the nature of research questions proposed here, it is quite possible no other approach could be suitable without significant alterations to original goal of this research project for a number of reasons.

As discussed above, the very formulation of the research questions proposed here would likely not happen employing any other approach: predictive capacity and behaviour modelling are inherently positivist notions, and are central to critical behaviourist approach. Any attempt to consider the subject matter employing any other alternative approach would inevitably require the modification of research questions, as research questions contemplated here are completely and profoundly interconnected with the quantitative scientific method of inquiry and with the underlying theoretical and philosophical aspects of positivist approach.

4.5 Research method

This section describes the research methods and design specifics employed: includes a comprehensive description of the sample, explains the research design, and describes statistical methods employed.

4.5.1 Sample

The Homescan data used here was acquired from the National Office of Statistics panel that comprises results from a survey of about 35,000 households and contains barcode scanned records of all their food purchases (was also used for example by Heravi & Morgan, 2014a; Heravi & Morgan, 2014b). The data arose from a market research data set supplied by Taylor Nelson Sofres (TNS), part of Kantar World Panel.

The subset selected for the analysis here covers only one product group: wine. This resulted in a subset with 170,989 cases, which cover purchases of 4,939 individual households over the time period from October 2002 through December 2005. A total of 224 variables are present in the dataset that include transactional, demographic, and product attributes – not all the attributes are usable however and many repeat variables are included (only usable for other product groups) which were omitted in the analyses here.

Data supplied by TNS UK Limited. The use of TNS UK Ltd data in this work does not imply the endorsement of TNS UK Ltd. in relation to the interpretation or analysis of the data. All errors and omissions remain the responsibility of the authors.

4.5.2 Research design

In previous work, informational and utilitarian reinforcement data acquired from matching studies (Foxall, Wells, Chang, & Oliveira-Castro, 2010) was integrated with the consumer behavioural data, effectively appending two additional variables to the dataset on a transaction level to reflect the informational and utilitarian reinforcers each brand was able to offer. As a result, for every case in the dataset that describes a brand purchasing decision, utilitarian and informational reinforcement parameters were used as independent variables (Greene, 2011). Even though it was shown that these additional reinforcement variables were able to contribute to the modelling, significantly improving the predictive capacity of the model, it is argued here that informational and utilitarian reinforcement as described in BPM (Foxall, 1990, 2004, 2005) are higher order latent variables that are formed using the regular variables such as product attributes and consumer demographics, and therefore ought to be represented by hidden layers in NN architecture.

4.5.2.1 Variables

Usually the process to identify the predictive variables to explain the relations with the dependent variable tend to be tedious and time-consuming: researchers start with a set of independent variables and identify the most predictive one, then begin to add more variables systematically and assess the model with R^2 or preferably adjusted R^2 and AIC value to decide whether it is worth adding the extra variables to improve the model. This process is very much a manual approach, and depends entirely on the perceptions of the researcher which variables to consider and in what order. There is also a matter concerning interactive variables, where interactive variables are sometimes produced and incorporated into the model – but even on those occasions only a few interactions between variables are considered. Once the interactions are considered, the choice of interactive variables to include into the model increases exponentially with the number of independent variables, and the manual process is unable to examine any kind of exhaustive list of interactions by any measure. This is one major drawback of the traditionally employed methods of analysis, as they all require predetermined structure specified during the modelling phase. Neural Networks on the other hand do not require any predetermined structure – connectionist models take complete input data and, during the training process, networks determine the best predictive variables and inherently examine all possible variables interactions, thus eliminating the otherwise necessary requirement to specify the network architecture beforehand. Various pruning methods are then able to strip the model further, producing the lean underlying structure that can serve as best estimation of underlying patterns within the data.

Price is an obvious choice for a dependent variable for any predictive modelling exercise, but from a purely semantic consideration, it becomes apparent that a straightforward price prediction may not provide a robust platform for the analytical work of explanatory nature. As previously discussed, *consumer choice* is a probabilistic value in behaviourist terms, and required to be operationalised as a proportion of instances of choosing one product over the other in a given time frame. Price alone could be overrepresented by the utilitarian reinforcement variables, as the majority of wine purchased by absolute volume or bottle count in the dataset used here would be a regular table wine. As such, the willingness of the consumer to pay more or less per litre of wine proposes a better modelling opportunity from the semantic and explanatory point of view as discussed in detail in the following chapters. Instead, considering the context of the BPM and theoretical framework adopted for this research, a new variable was generated as a ratio of expenditure to pack volume to be used as a dependent variable here: *price paid per litre*.

Connectionist models developed in such a way would then be useful in providing an insight into what underlying factors influence consumer choice situation and to what extent, and be able to assess quantitatively the changes to price per litre that the consumers are willing to pay based on the information available from independent variable values. Models could eventually contribute to the development of connectionist framework able to explain the consumer purchasing decision.

4.5.3 Apparatus

Statistical and data manipulation software employed during this research project include Microsoft Office Excel 2007 (Microsoft-Corporation, 2006), SPSS 17.0 (SPSS-Inc., 2007), R version 3.1.2 (R_Core_Team, 2014), and R studio (R_Studio, 2012). Initial compiling and appending of data was done in SPSS, which is most suitable to handle large data in addition to a reasonable degree of statistical analytical capacity – still inadequate for the connectionist modelling required here however should be noted. Excel was often used to quickly view the data due to its versatile nature – must be said it was completely useless for most of the statistical analyses however. R on the other hand is an exceptionally powerful application capable of performing advanced analyses – still, additional coding and package development was required to enable pruning and connectionist network visualisation. Developed by statisticians as an open source project, R consistently benefits from the contributions of worlds' leading analysts, which was used to developed and consequently assess all modelling work here. R Studio software merely packages R and provides a very user-friendly interface and additional functionality, which makes it much easier to use R.

4.5.3.1 R package development: RSNNS and NeuralNetTools

A large number of R packages were used, but the two most notable are the RSNNS (Bergmeir & Benítez, 2012) to do all the modelling and pruning, and NeuralNetTools to plot the connectionist network architecture once developed.

RSNNS was well developed to incorporate the functionality of the original Stuttgart Neural Network Simulator (SNNS) software (Zell et al., 1994; Zell, Mache, Sommer, & Korb, 1991a, 1991b, 1991c; Zell, Mache, Vogt, & Hüttel, 1993) into the R package as far as NNs modelling, but was lacking the pruning element – as did any other NNs package at the time. It was then essential to get involved with the original package developers, and introduce pruning as well which was implemented in the original software. For that, author performed the necessary coding in C++, which was then supplied to the package developers, who were able to code the necessary wrapper to integrate it into the R package. As a result, with a new update of the RSNNS, it now provides the pruning functionality to anybody who may be interested to pursue the connectionist modelling route as this author.

Somewhat similar events took place around the NeuralNetTools package, which was able to plot the connectionist architecture but not as far as the part when pruning occurred. After the collaboration of this author with NeuralNetTools package developers, it is now possible to plot pruned connectionist networks as well – a great way to represent visually the underlying architecture of the network as a result of model learning process.

4.5.4 Sequential account of the research process

This project builds upon previous work, and develops the subject further by addressing some of the limitations expressed heretofore (Greene, 2011). As such, the initial phase of this research project is consolidating the previous findings and identifying the line of inquiry that would be able to deepen the understanding of consumer behaviour. Extensive literature evaluation paired with ongoing discussions with specialists in the field of consumer behaviour is carried out to provide a comprehensive account of developments in theoretical framework of BPM as it is applied in empirical work. The connectionism is discussed in detail, giving particular attention to the philosophical developments and application of NNs models in consumer behaviour.

The second phase is largely revolving around the software development to enable the types of analyses and modelling work required here as described above. As a result, two R packages are improved to provide additional functionality.

This is followed by the next phase, which involves the models being developed. Some regression models are developed during the exploratory stages to learn the dataset, but their development is not progressed further as it is not within the scope of this research project – moreover, previous work already carried out the comparative analyses of logistic regression with NNs models (Greene, 2011). Instead, NNs models of varying architectural complexity are developed to examine the process of formulation of items in hidden layers of NNs model that could be interpreted as distributed representation of informational and utilitarian reinforcement as described in BPM.

Final phase built upon the modelling results acquired in the stages describe above, and encompasses a comprehensive discussion of implications on extending the BPM framework into the realm of connectionism. Future opportunities and research directions are identified, discussing the limitations of this research project.

4.6 Summary

This chapter focuses on unfolding the comprehensive account of research process in sequential manner to provide a thorough view as it is employed in this research project: an extended overview outlines the overall direction and goals, followed by a discussion on research questions, clarification of philosophical position, and justification and description of research methods with a sequential account of the process.

5. Analysis

This chapter discusses the statistical analyses employed, describes the specifics of the models developed in the course of research project, explains the tests in detail, and provides an overview of the results.

5.1 Preliminary data manipulations

The dataset was made available by TNS UK and the fieldwork was carried out as part of the Kantar World Panel and was originally obtained for the purposes of a different research study (Heravi & Morgan, 2014a, 2014b). The master database contains a large amount of data for a number of product categories: consumer household descriptive database contains all the household descriptors such as consumer demographics and household details, a transactional database contains all the purchasing data on individual transaction level, and product attributes database describes all the product SKUs in detail. The size of the database in MB was so large that it caused certain logistical problems while simply moving the files from one machine to another. As such, it was decided to focus on one product category – *wine* was selected for no other reason than author's previous experience in the wine industry, which involved working with wine data on a daily basis, thus offering a degree of familiarity with the data from the onset. Together with the colleagues from Cardiff Business School (Heravi & Morgan, 2014a, 2014b), the data was consolidated into a transaction level database, where household descriptors and product attributes were appended to the database. As already described above, this resulted in the initial database with 170 989 cases in total, and 319 variables. Even though the data contained an incredible number of descriptors, many of the attributes – product attributes in particular – were either not available or duplicated, no doubt used for other than wine product categories. Data manipulations initially were carried out on a superficial level only, without adjusting the values in any manner – only to ensure the data is easily transferable between different software applications employed here and would not cause any issues. During the exploratory analysis, many of the unusable variables – obvious duplicates and empty categories – were removed after initial examination, leaving the dataset with a total of 182 variables. Normally it would be beneficial to preserve the data as complete as possible, but the variables removed would not offer any contribution whatsoever and would rather create unnecessary noise and clutter.

5.2 Exploratory analysis

The compiled dataset contains a number of variables that describe individual households, product attributes, and more importantly the transactional data for every purchase occasion totalling at 170 989 cases: this makes it a transactional level data, as opposed to consumer level data where each case would represent individual consumer or household, most likely summarising the individual purchasing decision to some sort of an average or a sum value.

In the past author carried out similar research focusing on *consumer loyalty* as a dependent variable, which was operationalised with an additional variable calculated as a proportionate value of a money spent on the most often purchased product variant divided by the total amount spent on all product variants in a given time frame. This time however research questions deal with a decision-making process where the consumer makes a conscious decision to give up a certain amount of money in exchange for perceived utilitarian and informational reinforcement that the product – in our instance wine – would be able to provide. In particular, we are interested in the emergent process of assigning value to the unit of product, and therefore the operationalisation is as follows: dependent variable is the *price per litre* paid by the consumer, which is then predicted with different modelling methods using any consumer, product, and transactional independent variables. This new to the dataset dependent variable is calculated by dividing the total amount paid by a consumer household during each single transaction by a total number of litres of wine purchased, which produces a numeric monetary value. Important to note here that, unlike previous research that mainly focused on the predictive capacity of NNs models as opposed to traditionally employed methods such as logistic regression for comparative purposes (Greene, 2011), this research takes a step further and aims to model the emergent process and examine the explanatory capacity of NNs models in attempt to ultimately explain and even visualise the decision-making process.

5.3 Regression and NNs comparative analysis

To examine the relations in the data, some exploratory regression analyses were carried out.

To facilitate the modelling and avoid taxing the computational resources too much, a subset was selected for exploratory regression analyses by geography: all of Wales and West, which included 13787 total observations, 13782 once the few cases where price was 0 were removed. *Price per litre* variable was calculated here using the total *amount spent* divided by *volume*, and rounded to two digits (pennies). All transactions were then assigned to either *low spender* or *high spender* category, effectively converting *price per litre* into a new binary variable as follows: price 3.99 identified as a dividing point between the two very distinct types of purchases.

Variable conversion into binary is of course necessary for logistic regression analysis – it is a method of choice in extensive marketing research literature, and has been proven to offer consistent level of insight when it comes to modelling relations in the data (Adya & Collopy, 1998). Starting with a logistic regression establishes a solid basis for the ongoing analysis, and R (R_Core_Team, 2014) offers ample solutions with multitude of packages that provide the logistic regression functionality to examine the data.

Zelig (as described in Crosas, King, Honaker, & Sweeney, 2015) is one useful package in R to build regression models, and is used here for comparative purposes. Using only eight independent variables to predict the binary high or low spender dependent variable, a regular least squares regression model is developed with the following results: multiple R-squared of 0.07001, and adjusted R-squared of 0.06947. Rather low R-squared values are expected however while trying to predict what type of spender consumer may be from a few simple descriptive attributes, and may indicate a challenge regression model faces using the consumer behaviour dataset working with the dependent variable converted from the probabilistic to binary. Price per litre is a numeric monetary value, and when converted to either low or high value binary in nature is bound to have an effect on R-squared values. Low R-squared values may also be due to the fact that relevant variables are not available or were not measured in a suitable manner, or that the model is not able to account for other effects, such as nonlinearity. In validating models of consumer behaviour however, the R-squared value may not be as important as other measures: within-sample prediction accuracy for instance, or the coefficient values and the properties of the consumer response promptness to price changes. As this research is not particularly concerned with the R-squared value itself, but rather with its determinants and underlying structural effects, the inherent nonlinear effects of NN models are expected to provide quite the improvement with connectionist models discussed later.

The data is then randomly split into two subsets for validation – *regression-NNs validation subset 1* and *regression-NNs validation subset 2* – and models are developed using both subsets independently in parallel. As a next step, Logit regression models are developed with Zelig package for each of the subsets using the same variables as with least squares regression described above, and all variables are statistically significant as shown in Figure 4 using *regression NNs validation subset 1*.

```
Coefficients:
             Estimate Std. Error z value Pr(>|z|)    
(Intercept)  0.34297    0.09002   3.810 0.000139 ***
AdultNo_    -0.18625    0.03811  -4.887 1.02e-06 ***
ChildNo_    -0.24651    0.03734  -6.602 4.07e-11 ***
Diabetics   -0.11975    0.02935  -4.080 4.51e-05 ***
Tvhours     -0.07448    0.01659  -4.489 7.15e-06 ***
Cars         0.54424    0.03629  14.999  < 2e-16 ***
Cats        -0.15037    0.03001  -5.010 5.44e-07 ***
Dogs        -0.34581    0.04507  -7.673 1.68e-14 ***
TVs         -0.10172    0.02247  -4.527 5.97e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 4. Regression coefficients and significance levels for *regression-NNs validation subset 1*.

Even though linear regression models and NNs parallel connectionist models are fundamentally different, it is beneficial nevertheless to try to bring them to the same level of analysis as a benchmark and compare the regression variable contribution parameters that the logistic regression model provides with the weights of the NNs model. Connectionist models are very powerful algorithms for a number of reasons – not the least is inherent parallel nonlinear configuration that requires no predetermined model structure. For the purposes of comparative exercise, it is possible to isolate this nonlinear capacity and limit the NNs models to only 2 layers of nodes, effectively constraining the model to a linear function: input and output layers. Thus, using the same variables, the simplest 2-layer NNs model is developed using a simple and elegant *nnet* package in R that offers functionality to satisfy the NNs modelling research in most cases

(Venables & Ripley, 2002) – no hidden layers, just two layers for input and output nodes. As a result, the 8-0-1 NNs model contains 8 weights as shown in Figure 5.

```
a 8-0-1 network with 9 weights
options were - skip-layer connections  entropy fitting
 b->o i1->o i2->o i3->o i4->o i5->o i6->o i7->o i8->o
 0.34 -0.19 -0.25 -0.12 -0.07  0.54 -0.15 -0.35 -0.10
```

Figure 5. NNs model results for *regression-NNs validation subset 1*.

What immediately becomes obvious once the regression and NNs results are examined is that NNs mode weights are identical to regression coefficient values – the simplest NNs model with no hidden layers performs in exactly the manner as logistic regression.

This is confirmed by using *regression NNs validation subset 2* as well, and NNs model weights are again identical to those of logistic regression coefficients as shown in Figure 6 and Figure 7.

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)    
(Intercept)  0.28470    0.09170   3.105  0.00191 ** 
AdultNo_    -0.12489    0.03878  -3.221  0.00128 ** 
ChildNo_    -0.25516    0.03747  -6.809 9.80e-12 ***
Diabetics   -0.05688    0.02785  -2.043  0.04109 *  
TVhours     -0.07994    0.01673  -4.779 1.76e-06 ***
Cars         0.51149    0.03647  14.025  < 2e-16 ***
Cats        -0.12054    0.02999  -4.020 5.83e-05 ***
Dogs        -0.35629    0.04612  -7.726 1.11e-14 ***
TVs         -0.11419    0.02285  -4.998 5.81e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 6. Regression coefficients and significance levels for *regression-NNs validation subset 2*.

This demonstrates how a simplistic NNs model with no parallel nonlinear processing performs exactly as logistic regression would, and provides connection weights identical to regression coefficient values.

```
a 8-0-1 network with 9 weights
options were - skip-layer connections  entropy fitting
 b->o i1->o i2->o i3->o i4->o i5->o i6->o i7->o i8->o
 0.28 -0.12 -0.26 -0.06 -0.08  0.51 -0.12 -0.36 -0.11
```

Figure 7. NNs model results for regression-NNs validation subset 2.

To increase the validity and reliability of the test, this procedure was replicated 1000 times: random split into 2 new subsets each time which are then used to build logit and NNs models to compare the regression coefficient values with NNs connection weights, providing analogous results every time.

5.4 Exploratory NNs modelling

Following the exploratory analysis with traditional methods such as regressions, initial NNs modelling was carried out. This phase was mainly concerned with testing the software capacity to select a suitable package in R to carry out the modelling. Upon examination of the common R packages that offer NNs modelling functionality, it became apparent that only a few offered the capacity to use multiple hidden layers, and – more importantly – none offered the capacity to carry out pruning. This meant that author was facing a few options: either abandon R in preference of another software, develop a new package from scratch, or work with one of the existing package developers to advance the functionality and expand it to include pruning. First option was easily dismissed, as R is one of the most advances statistical modelling platforms: if something was not available in R, very likely this was not available in other software packages either and would have taken significantly longer to develop and roll out to make it available than it would take to do the same in R, as the statistical programming environment for R is structured around developing new packages by researchers for other researchers. Second option implied the authors would be highly proficient with coding, which was not the case and developing the skill would require significant amount of time and could be considered a research project in its own right. Thus, the third option was pursued where collaborating with coders and developers additional functionality for the existing R package was developed to allow pruning the networks.

5.4.1 RSNNS

RSNNS (Bergmeir & Benítez, 2012) was an R package that allowed multiple hidden layers, and Bergmeir, package developer, was very responsive to the initial communication and general questions regarding the R package. This author proposed a collaboration project to develop the R package to include pruning, and Bergmeir offered assistance with wrapping the R code and updating the package once the coding is complete. The underlying functionality and low-level interface was done in C++, which meant this author had to develop a sufficient enough level of understanding and skill to do the necessary coding. This alternative however was the most feasible and an optimal choice.

Package development began with familiarising with the low-level interface, which was based on C++ coding and already available for NNs training functionality, but not pruning. Using the *SNNS* manual and the ad hoc assistance of package developer that involved adding missing low-level functions to the package, with considerable effort this author was able to compile initial code capable of carrying out pruning using the low-level interface, which was then implemented and updated by Bergmeir in the official package and made available to anybody within the wide scientific community.

As *nnet* package only allows connectionist networks with a single hidden layer and no pruning capability, *RSNNS* package is used for all consecutive connectionist modelling from here on.

5.4.2 NeuralNetTools

NeuralNetTools (Beck, 2015) is a rather recent package used to visualise the NNs architecture. When it was first encountered, this author found the flexibility in visualisation for a number of NNs R packages extremely useful, including package *nnet* described above, which was also used extensively in previous research (Greene, 2011), and package *RSNNS* as well. Even though initially *NeuralNetTools* was great for visualisation of the NNs architecture, it was unable to accommodate the pruning process. When this author contacted the *NeuralNetTools*' package developer Beck with an inquiry, he was not aware pruning was in fact available at all in R – not surprisingly so, as pruning in *RSNNS* was only introduced from this author's collaboration with *RSNNS'* Bergmeir (2012) very recently. *NeuralNetTools'* Beck was happy to advance and develop the visualisation capacity further, and after collaboration with this author, *NeuralNetTools* had a new functionality that allows visualising pruned NNs model architectures as well.

5.5 Advanced connectionist models: hidden layers

The simplistic NNs model described above did not use any hidden layers and therefore did not develop any capacity to account for nonlinearity in the architecture. Next, hidden layers are introduced in the connectionist modelling to provide a nonlinear dimension and build advanced models of consumer behaviour. For illustration purposes, number of variables used in the model is increased as reflected in a more complex input structure, and a single hidden node is introduced to the network architecture as shown in Figure 8. Black connections identify reinforcing connections, whereas light grey connections identify inhibiting connections – line weight corresponds with the connection weight. A single node in the hidden layer would not be expected to contribute to the explanation as compared with what a regression would normally be able to offer however, so the number of nodes is gradually increased.

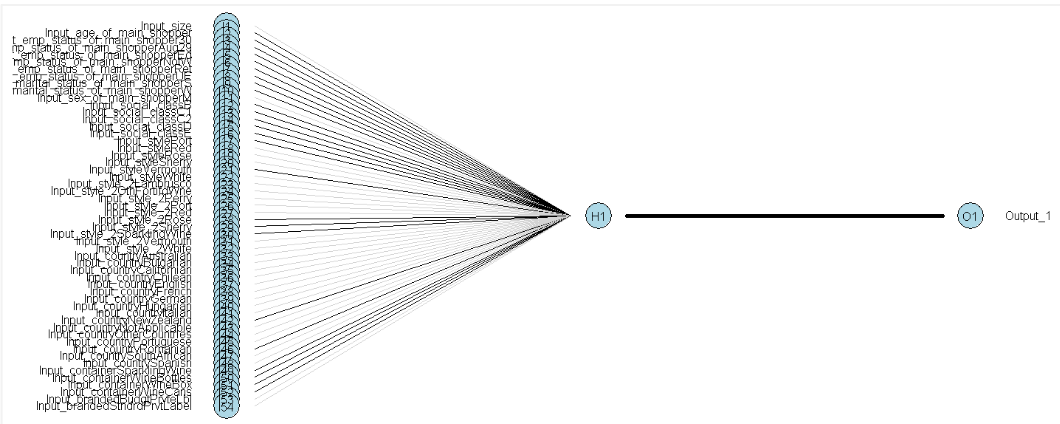

Figure 8. Connectionist network 54-1-1 architecture using consumer data with 1 hidden layer and a single neuron, 1000 iterations, no pruning.

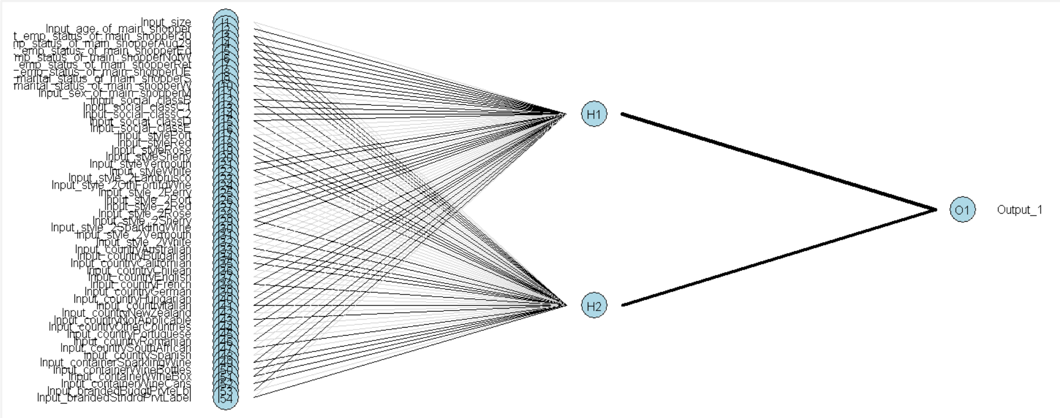

Figure 9. Connectionist network 54-2-1 architecture using consumer data with a single hidden layer and 2 neurons, 1000 iterations, no pruning.

Figure 9 shows 2 nodes in the hidden layer: simply increasing a number of nodes to a total of 2 results in a dramatic change in the network architecture, as now the hidden layer is able to examine the nonlinear relations within the data.

If the number of hidden nodes is increased to 4 as shown in Figure 10, the architecture becomes even more complex, but at the same time provides the network with an additional capacity to extract microfeatures.

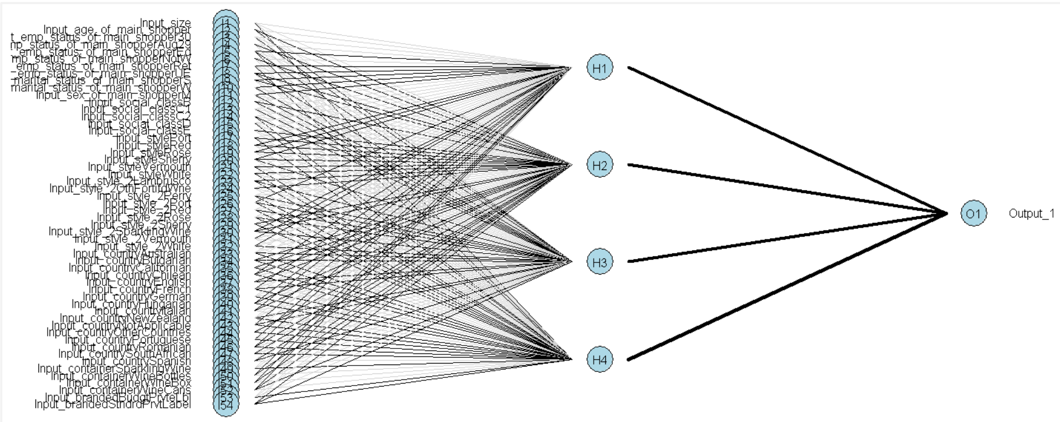

Figure 10. Connectionist network 54-4-1 architecture using consumer data with a single hidden layer and 4 neurons, 1000 iterations, no pruning.

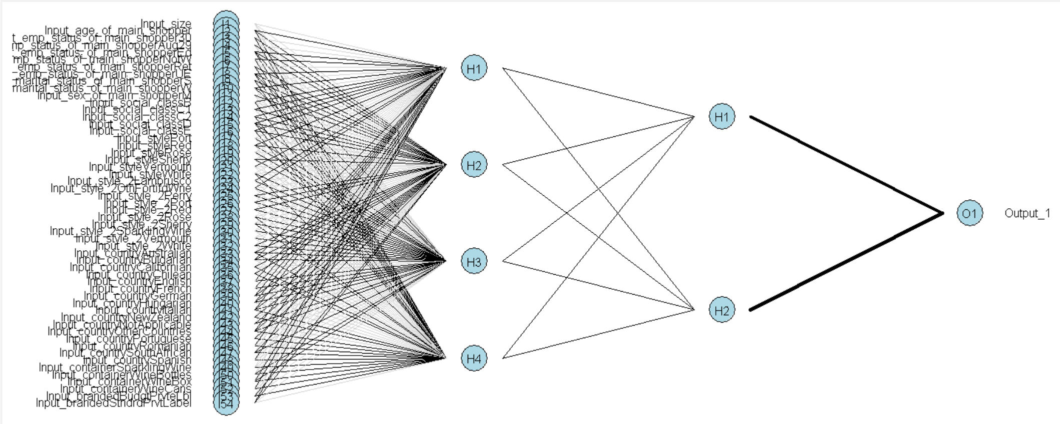

Figure 11. Connectionist network 54-4-2-1 architecture using consumer data with 2 hidden layers and 4-2 hidden neurons, 1000 iterations, no pruning.

As discussed above, in addition to increasing the number of hidden neurons within a single hidden layer, another way to increase the connectionist model complexity is by increasing a number of hidden layers and distributing the hidden neurons among multiple layers. Figure 11 shows connectionist network with the 54-4-2-1 structure, where a total of 6 hidden neurons are distributed among 2 hidden layers, and Figure 12 shows the network that incorporates yet another hidden layer for the total of 3, with 14 hidden neurons in a 54-8-4-2-1 network structure. This of course provides an innumerable number of options how the initial network architecture could be arranges – in the following sections some of this will be examined and explored in attempt to assess which type of network architecture could be more suitable for either predictive or explanatory purposes.

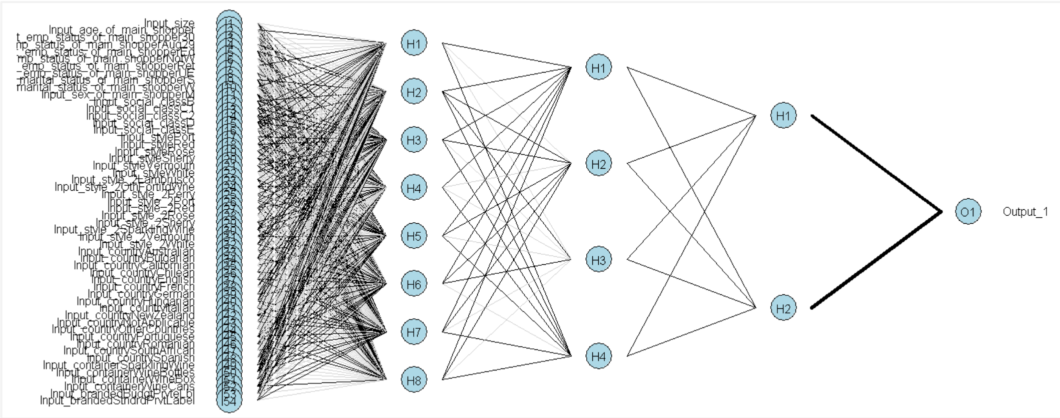

Figure 12. Connectionist network 54-8-4-2-1 architecture using consumer data with 3 hidden layers and 8-4-2 hidden neurons, 1000 iterations, no pruning.

It is of course not the complexity in itself that we are after here, but rather the sufficient flexibility and functional size of the initial network architecture to allow the further developments such as pruning to be carried out in the best possible manner, and provide optimal result. Nevertheless, this capacity to account for such a level of complexity is what makes the connectionist networks such a powerful pattern recognition algorithm as it allows extracting microfeatures that may very well even be incomprehensible to human researchers, which at the same time makes the interpretation extremely difficult or perhaps even impossible. A number of variable contribution analysis methods exist that attempt to examine the network connection weights and interpret the results will be discussed later, but here the focus is on pruning methods as a preferred method for the connectionist network to systematically eliminate some connections that do not contribute to the explanation.

5.6 Pruning

The dataset used here is still the same subset for Wales and West as above, but only 1024 cases are randomly selected to test the models on initial development stage. Reason for a somewhat conservative number of cases is that it takes quite a long time to process the model computationally: could be only a few seconds for a simple model as above with 1000 iterations and a couple of nodes in the hidden layer, or considerably longer for a for a large model with 3 hidden layers and 100 000 iterations with retrain pruning cycles – as long as several days of non-stop processing.

Before advancing to modelling with consumer data however, it would be worth to examine and assess pruning process with simulated data and tasks – i.e. train a NNs model using more weights than the task requires, and examine how well pruning deals with eliminating the unnecessary parts to trim down the model and expose the underlying core architecture to explain relations in the data.

5.6.1 Assessing pruning performance using simulated data

Before proceeding with modelling consumer data, it is important to assess the pruning capacity to isolate and remove unnecessary connections. It would not be feasible to test this with consumer data in which the relations and patters are not yet established – in fact, this is something this research project aims to achieve to an extent. Thus, simulated data would be used.

Simulated dataset would contain *X1* and *X2* values, which are randomly generated figures between 0 and 1, 1000 items each. *Y* would be the following function with some noise added in R:

y = sin (x1 - 2 * x2) + 1 + rnorm (1000, 0, 0.2)

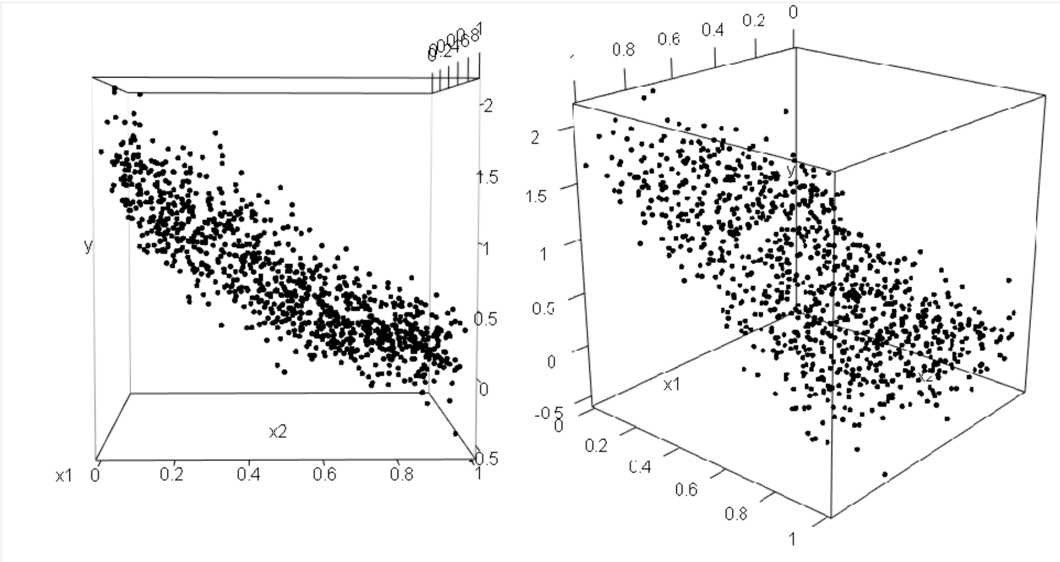

Figure 13. Visual representation of y = sin (x1 - 2 * x2) + 1 + rnorm (1000, 0, 0.2)

In Figure 13 the data is visualised in a 3-dimensional plot, showing that the cases may seem to form a linear relationship in a 2-dimensional a surface, whereas a 3-dimensional representations shows the data to be organised around a surface rather than a line – to solve this, the connectionist model would require at least one hidden layer with just 2 nodes, making all extra nodes unnecessary and redundant. If that is indeed the case, pruning algorithm should eliminate the unnecessary superfluous architecture, leaving the bare minimum core necessary to solve the problem: 2 nodes in a single hidden layer. To test this, a simplified connectionist model with excessive 2-2-2-2-2-1 architecture is developed – that is a connectionist network with 4 hidden layers in addition to input and output layers containing 2 nodes within each of the hidden layers. As suggested above however, this problem only requires a single hidden layer with 2 nodes, thus pruning should strip out the rest of the network exposing the core network architecture.

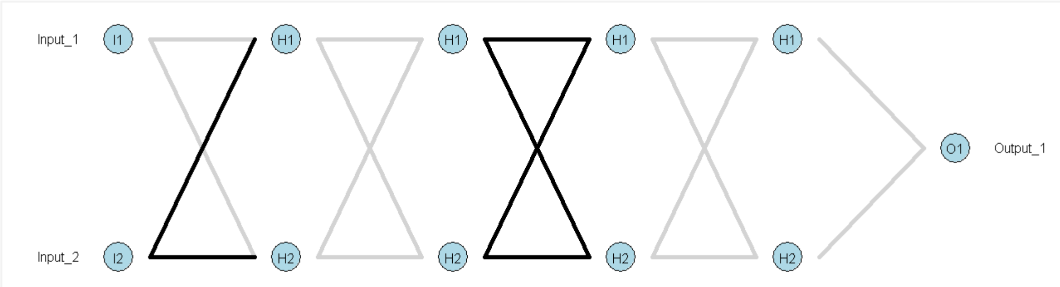

Figure 14. Connectionist model with 2-2-2-2-2-1 architecture design, no pruning used in model development.

No pruning was used to build the network shown in Figure 14 – as a result, it is a fully connected network that uses all the hidden layers and nodes, even though it would not be required to solve this particular computational problem. This illustrates the issue with predefined model architecture as it is difficult and even impossible to know what exactly would be required to solve the task with consumer data before the data is actually examined – a case of circular logic even. With the synthetic artificial data used here it is well known what is required to solve the task, and are able to say definitively that the predefined network proposed here for illustrative purposes that incorporates excessive and redundant architecture is able to solve the task just fine, but at the same time makes it particularly difficult to examine the network architecture in attempt to explain and define relations between the variables and interpret the results.

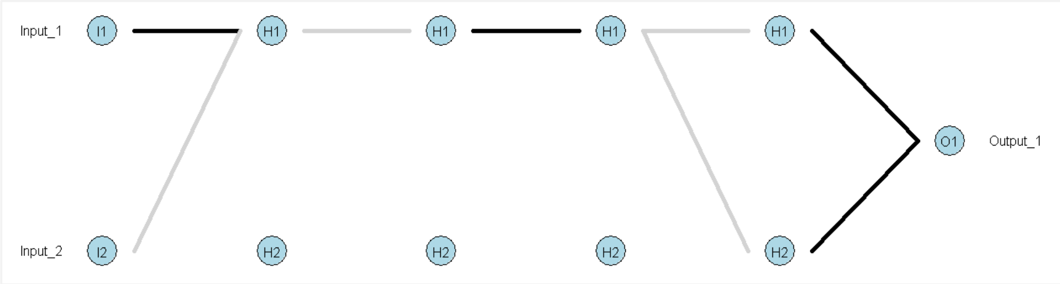

Figure 15. Connectionist model with 2-2-2-2-2-1 architecture design, pruning used in model development.

For the model shown in Figure 15 exactly the same steps are taken to develop the connectionist model using the same synthetic data as for the model shown in Figure 14 above, but now employing pruning methods to optimise the network architecture during the model learning process. *Optimal Brain Surgeon* algorithm displayed better results with consumer data when all available in RSNNS pruning methods were tested, and is used for pruning here (Hassibi & Stork, 1993). Using *Optimal Brain Surgeon* algorithm offers substantial benefits for connectionist models: improved generalisation, simplified network architecture, reduced computational capacity required and improved processing time as a result, and – crucial for the research questions postulated here – improves network rule extraction capacity as a result. It now becomes apparent that the core architecture necessary to solve the task is in fact much simpler than shown in Figure 14 and indeed only requires a single hidden layer and 2 nodes – extra neurons within the first, second, and third hidden layers are all redundant now, as the 2 neurons within the last hidden layer are sufficient to solve the computational problem. The level of architecture complexity required for the task here overall is rather low, as we only have 2 nodes in the input layer and few hidden nodes – yet it is apparent that the pruned network architecture is substantially more clear and easier to examine and interpret as a result. Pruning a highly complex network architecture that would be trimmed to the core network as a result should, potentially removing multiple nodes, should provide a substantial benefit in explaining network architecture.

Since the synthetic dataset was designed for a network that would only require a single hidden layer with only 2 nodes, we could assess performance of the 2 models above and compare it precisely with that model architecture: a simple model with only 2 hidden neurons within a single layer. As a result, the network architecture is as shown in Figure 16: fully connected with 2 hidden nodes. Effectively, this architecture is similar to the one described above after the pruning was carries out (shown in Figure 15) as pruning removed all but the 2 hidden nodes in the last layer only.

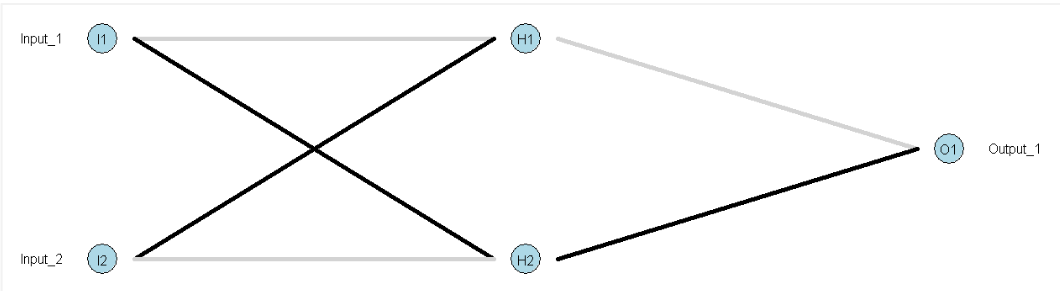

Figure 16. NNs model with 2-2-1 nodes in a single hidden layer, no pruning.

It would be useful to compare model performance of all 3 models described above: 2-2-2-2-2-1 model with no pruning (Figure 14), 2-2-2-2-2-1 model with pruning carried out (Figure 15), and then 2-2-1 model with a single hidden layer (Figure 16).

Using RMSE, pruned 2-2-2-2-2-1 models (Figure 15) show RMSE output around 0.28, similar to 2-2-1 models with a single hidden layer and no pruning RMSE output around 0.27 (Figure 16) – interestingly enough, 2-2-2-2-2-1 models with no pruning (Figure 14) provide RMSE that varies between higher values around 0.53 and similar to other models 0.27. This procedure was replicated 1000 times, providing consistent results.

This should serve as a convincing argument that pruning connectionist models is an optimal approach to develop robust representative models to extract and describe representative patterns within the data, and would be exceptionally useful to examine and potentially explain consumer behaviour and decision-making process of consumer choice. Moreover, the tests carried out suggest that pruning is not ably able to substantially simplify the network architecture and reduce the network size by eliminating the inessential connections, but able to do so while the maintaining the level network predictive performance on a similar or better level as compared to models where no pruning was carried out.

5.6.2 Assessing pruning performance with consumer data

Now that it is established that pruning is an effective and efficient way to remove unessential connections to extract relevant patterns in the data and reveal the core relations that may be able to explain consumer behaviour and decision-making process, it would make sense to proceed with consumer data modelling.

Using the same Wales and West consumer data sample subset as in the section describing comparative analyses above but with additional input variables, the connectionist modelling is carried out to develop a predictive and explanatory representation of consumer behaviour. For the first set of exploratory models, only 1024 randomly selected cases are used – this is of course to speed up the modelling process where a number of preliminary starting model architectures are examined. Even with a method of analysis that does not require the final model architecture to be predetermined and thusly to a large extent defined by a researcher, there is still a matter of initial model architecture which needs to be defined by the researcher nevertheless – i.e. the number of hidden layers and computational nodes to be used and consequently pruned. This will be discussed in detail in the limitations section later.

Number of iterations (and number of retrain cycles with pruning) plays an important role as well: using 3 hidden layers with 8 neurones within each layer it takes 10.15 sec with 1000 iterations, this goes up to 63.40 sec with 10 000 iterations, and up to 637.25 sec with 100 000 iterations. This network is depicted in Figure 17 where the 54-8-8-8-1 structure contains an input layer with 54 nodes (some of the non-numeric inputs are automatically dummy coded and therefore appear as separate input nodes here), 3 hidden layers, and an output layer can be clearly seen. Even though the connection weights are clearly identified by the connection line weight, it is nevertheless very difficult to comprehend such a complicated network visualisation displaying 256 connections between 79 neurons, making any attempts at interpretation problematic. This is not to say that the network shown here is extremely complex – in earlier research, substantially higher numbers of neurons were examined to assess the predictive capacity in relation to connectionist model size, going as high as 200 neurons within a single layer. Thus, it is possible that a more complex starting network architecture would be required here to eliminate the possibility of inadvertently limiting the final network architecture by setting too small of a starting point – this of course assuming that pruning algorithm should consequently be able to eliminate any number of redundant connections and neurons within the overall architecture.

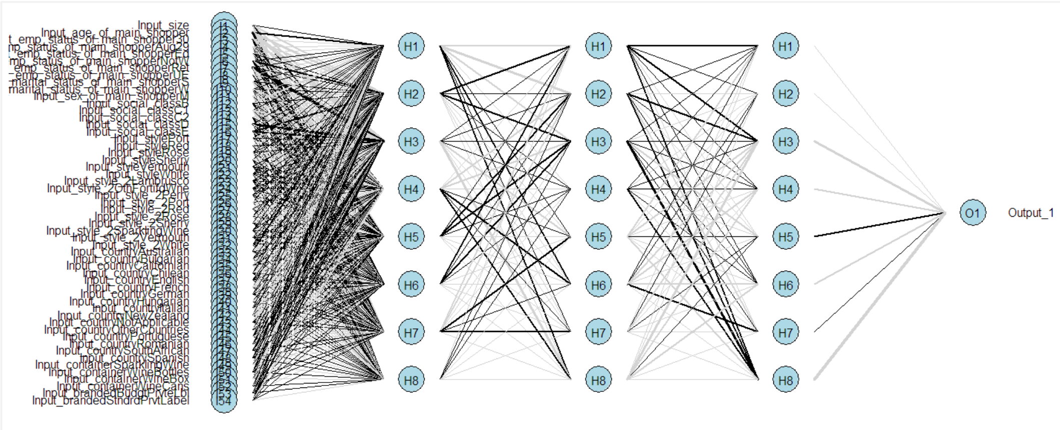

Figure 17. Connectionist network 54-8-8-8-1 architecture using consumer data with 3 hidden layers and 8 neurons each, 100 000 iterations, no pruning.

This is a rather straightforward network architecture as far as number of hidden layers and neurons – as a next logical step, it would be useful to introduce pruning and examine the effects. Figure 18 shows a model described above: three hidden layers with 54-8-8-8-1 architecture, using 1000 iterations and no pruning: it takes only 10.15 sec to complete and produces 568 connections in total with 79 neurons. Once pruning is introduced with 1000 iterations and only 100 retraining cycles, it takes considerably longer to calculate: 306.25 sec. The resulting network architecture however makes it abundantly clear how helpful the pruning process really is even from just looking at the network in Figure 19 as compared to network architecture in Figure 18 where no pruning was done – a large number of connections is pruned out decreasing the total number to just

85, a number of neurones are also pruned out leaving a total of 74, and the connection weights appear to be higher.

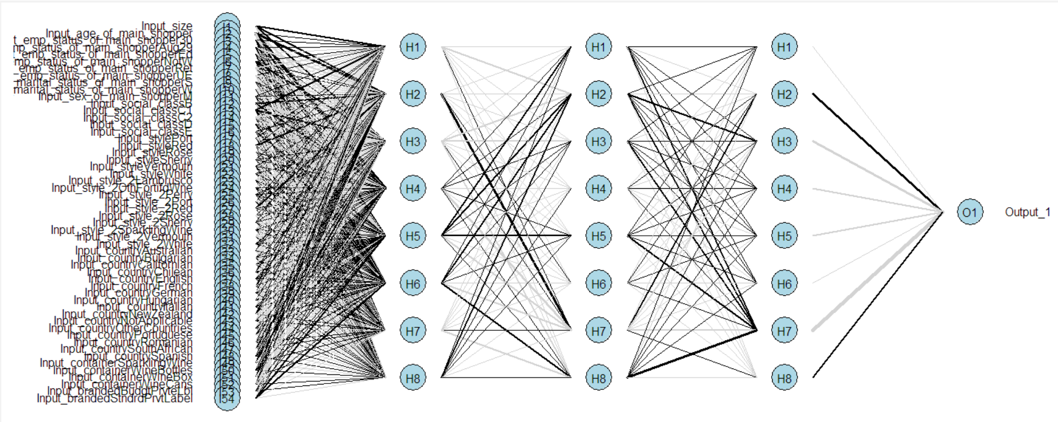

Figure 18. Connectionist network 54-8-8-8-1 architecture using consumer data with 3 hidden layers and 8 neurons each, 1000 iterations, no pruning.

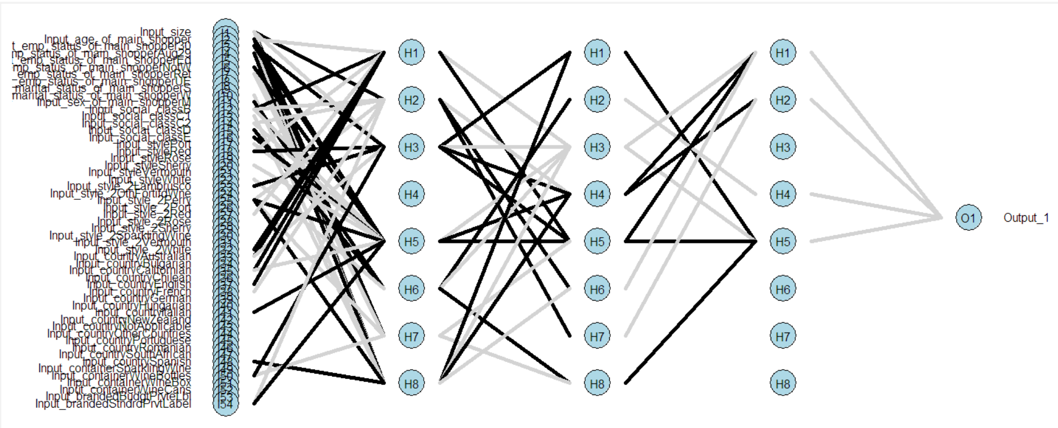

Figure 19. Connectionist network 54-8-8-8-1 architecture using consumer data with 3 hidden layers and 8 neurons each, 1000 iterations, pruning with 100 retrain cycles.

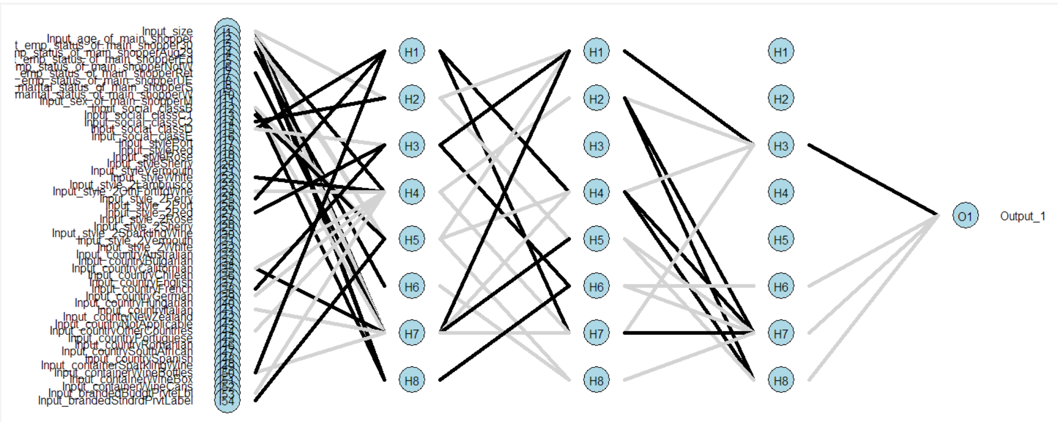

Figure 20. Connectionist network 54-8-8-8-1 architecture using consumer data with 3 hidden layers and 8 neurons each, 1000 iterations, pruning with 250 retrain cycles.

With 250 retraining cycles as shown in Figure 20, it now takes 736.28 sec to complete, and the network architecture is further optimised down to 66 connections in total with 74 neurons.

It is clear from these few examples the general direction the modelling process should take, and suggests that pruning could be of great help in reducing network architecture complexity to improve the simplify the data interpretation for explanatory purposes.

5.6.3 Pruning different network architectures

As described above, the connectionist network determines and defines the final network architecture during and as a result of the learning and pruning process. Researcher however does need to define the initial design of connectionist network: number of hidden layers and number of computational neurons within each hidden layer are set before the model learning process begins. Given the nature of research questions this research project is mainly concerned with, it is important to examine the effectiveness of pruning different connectionist network architectures. It makes sense to proceed with a sufficiently large initial network architecture not to limit the model capacity from the onset, but a compromise is necessary between including a larger dataset than used in previous tests described above and keeping the modelling time reasonable. Thus, the overall network size will be limited to 12 computational units here – this will keep the network size reasonably constant and will allow focusing on manipulating network architecture. There are a number of network designs to allocate hidden neurons within 1-, 2-, and 3- layer networks, which will be discussed in the following paragraphs.

To begin with, the network architecture with a balanced 54-4-4-4-1 layout is examined. Using a larger consumer data subset this time that includes all of Wales and West regions with a total number of 13787 cases, it takes 503.84 sec to complete and produces a fully connected 54-4-4-4-1 network with 67 neurons linked by 252 connections as shown in Figure 21, and produces RMSE of 0.9087.

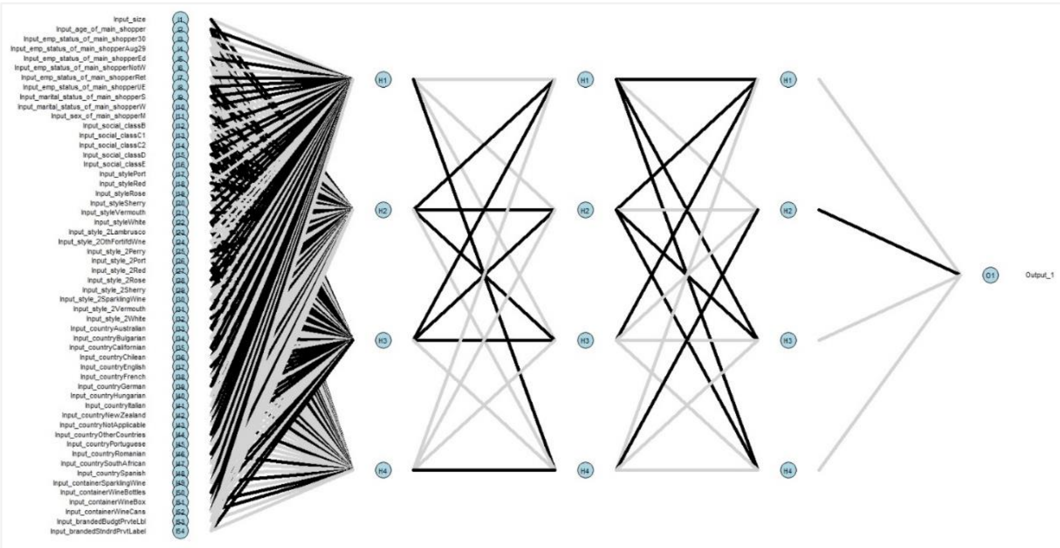

Figure 21. Connectionist network 54-4-4-4-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 10000 iterations, no pruning.

Using the same data and 54-4-4-4-1 network structure, the connectionist network is developed applying pruning with 100 relearning cycles. It now takes quite a bit longer to train and prune the network – a total of 1709.36 sec – but as a result produces a network with much leaner architecture, only 59 neurons and 56 connections, and RMSE of 0.9538, which is comparable to the unpruned network shown in Figure 21. In fact, as can be seen in Figure 22 where the network architecture is shown, the pruning was successfully able to effectively remove a few hidden layers leaving only 2 nodes in the first hidden layer. The network architecture shown in Figure 22 is the outcome of one of the first modelling attempts, and indeed the result is extremely promising as far as theoretical implications go, and what it may mean for the discussion of using connectionist models to substantiate and extend the theoretical framework of BPM.

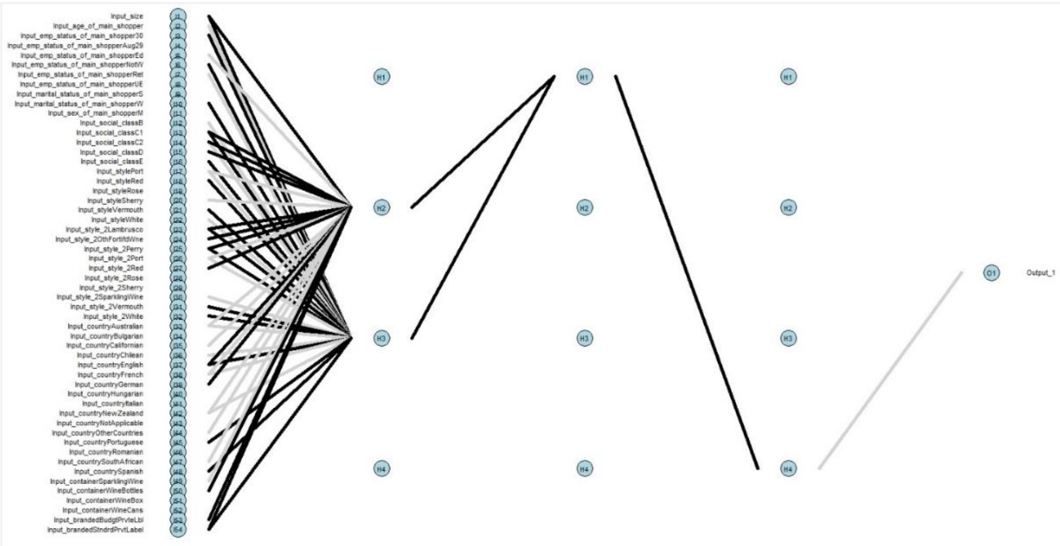

Figure 22. Connectionist network 54-4-4-4-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 10000 iterations, pruning with 100 retrain cycles.

This of course needs to be validated, with the procedure replicated numerous times to see if consecutive models continue providing consistent results and similar patterns in the network architecture can be observed.

To examine the changes in the network architecture as a result of pruning, a number of different architecture types are examined here for exploratory purposes. Using the same number of neurons within the network as in the model shown in Figure 22, 3 different network architecture types are explored: 54-6-4-2-1 network that funnels the connections through the hidden layers, 54-4-4-4-1 network which analogous to the one shown in Figure 22 used for validation and benchmarking, and reverse funnel 54-2-4-6-1 network that imposes a bottleneck within the first hidden layer but allows growing the network slightly throughout the successive hidden layers. Each of these network types was developed as a trained (no pruning) and pruned variant, and replicated 100 times with 10000 iterations and 100 retrain cycles (*Optimal Brain Surgeon* pruning algorithm shows remarkable performance with pruning and does not actually require many retraining cycles) for validation purposes, producing a total of 600 models as a result. Figure 23 shows one of the most common architectures as a result of pruning the 54-4-4-4-1 network, which was optimised from 67 units with 252 connections down to 64 units with only 67 connections. It is clear that the network effectively eliminated certain neurons altogether by pruning their connections.

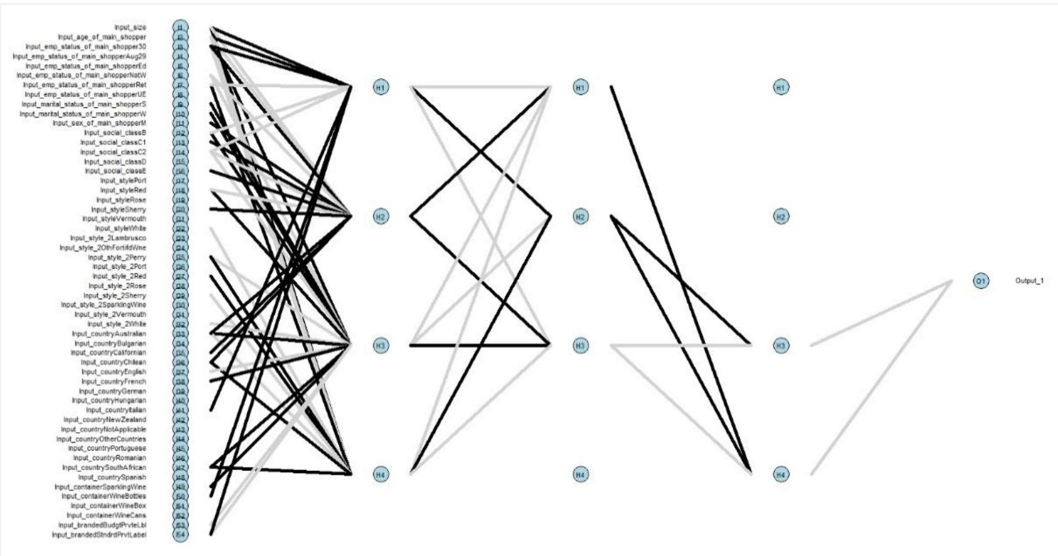

Figure 23. Connectionist network 54-4-4-4-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 10000 iterations, pruning with 100 retrain cycles.

Due to the nature of the data employed here with a large number of input neurons, the initial architecture of the network with 54-6-4-2-1 contains substantially more connections (between input and first hidden layer) – as a result, the common network architecture type obtained here was optimised from 67 units with 358 connections down to 67 units with only 78 connections as shown in Figure 24.

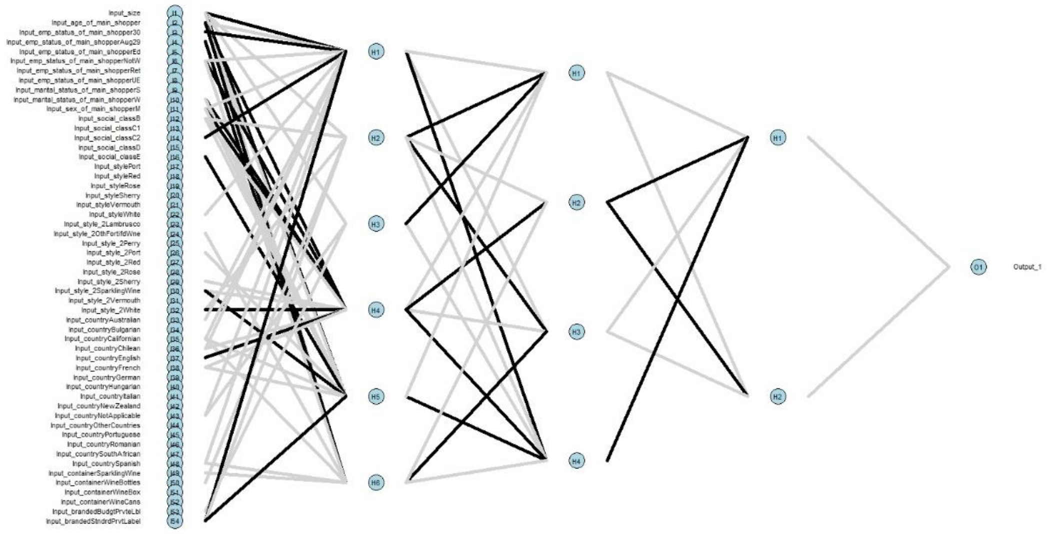

Figure 24. Connectionist network 54-6-4-2-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 10000 iterations, pruning with 100 retrain cycles.

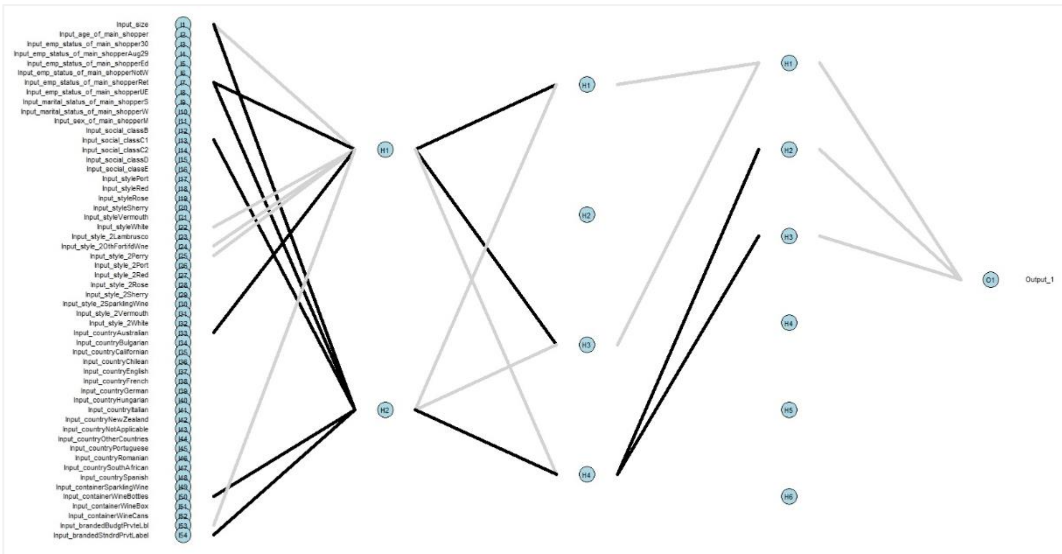

Figure 25. Connectionist network 54-2-4-6-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 10000 iterations, pruning with 100 retrain cycles.

Models with initial 54-2-4-6-1 network architecture design contain 67 units and 146 connections, and can be pruned down to 63 units with only 25 connections, as shown in Figure 25. Here, the bottlenecked network did not show the capacity to sufficiently develop within the subsequent hidden layers, normally leaving a number of hidden neurons unused.

These results support the premise that even the smaller neural network models should be able to extract the highly complex patters from the consumer decision-making data and represent it with a visually straightforward network architecture, while being able to carry a substantial amount of predictive and explanatory capacity. If this is indeed the case, a simplified version of the network architecture with a few hidden nodes should serve as a great explanatory model of consumer behaviour – say a network with a 54-2-1 architecture as discussed in the following section.

5.6.4 Assessing explanatory capacity of a connectionist model with a single hidden layer and 2 neurons

The simple approach would be to include only 2 hidden nodes – this way, the model is able to account for nonlinear relations within the data, and all possible interactive combinations of input nodes would be accounted for and linked to the output node through the 2 hidden nodes. It then should be possible to argue that the 2 hidden nodes in this simple connectionist architecture represent utilitarian reinforcement and informational reinforcement following the BPM framework. Figure 26 shows this 54-2-1 network, where only 2 hidden nodes are used with 1000 iterations – very simple straightforward network that only took 5.69 sec to run and uses 57 nodes (54 of which are input nodes as can be seen in Figure 26) and 110 connections.

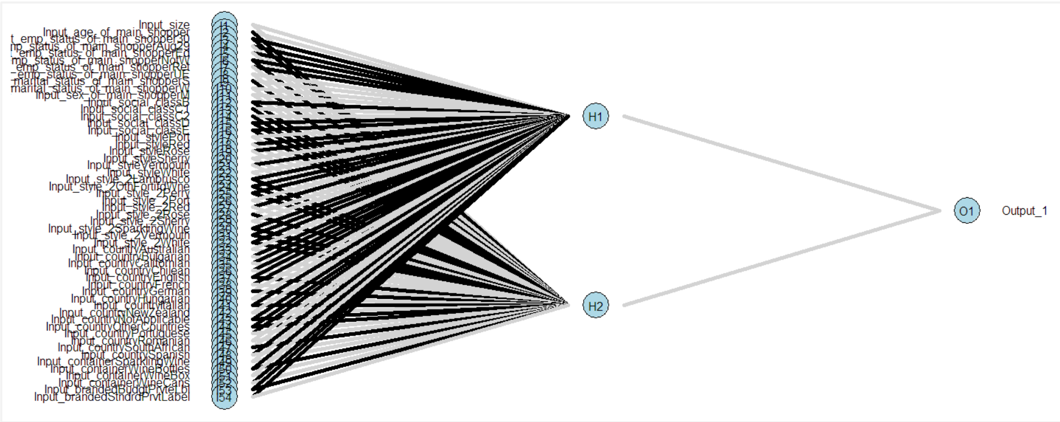

Figure 26. Connectionist network 54-2-1 architecture using consumer data with a single hidden layer and 2 neurons, 1000 iterations, no pruning.

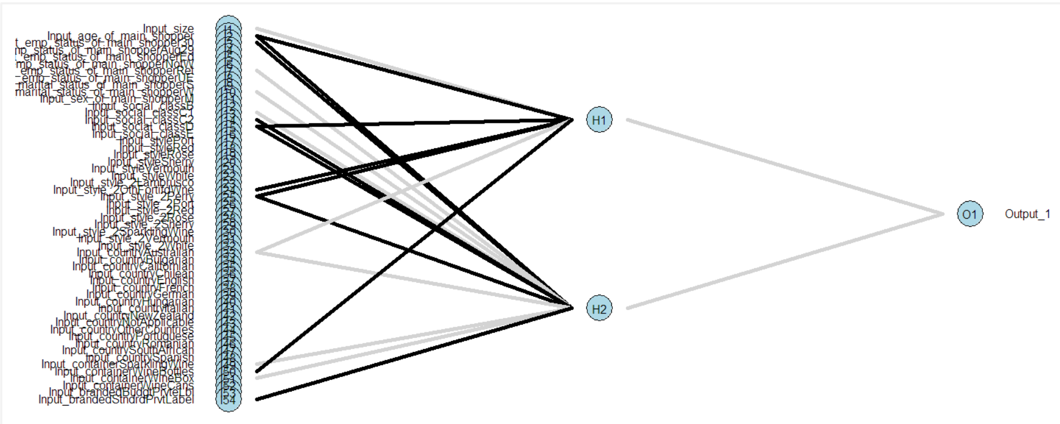

Figure 27. Connectionist network 54-2-1 architecture using consumer data with a single hidden layer and 2 neurons, 1000 iterations, pruning with 100 retrain cycles.

Once pruning is introduced in the network shown in Figure 27 with 100 retraining cycles – it takes 20.04 sec to run and pruning algorithms removes all but 21 connections as a result. Obviously, this is a very straightforward and easy to interpret – any one of the input nodes can be traced to one of the hidden nodes, and to output – not only visually, but also quantitatively, as every connection carries the connection weight of course. This should allow the examination of the architecture in detail and assess the variable contribution level. The interpretation of result should be very transparent as well, where all connections are represented visually.

However, would it allow sufficiently robust network to develop before it is pruned? Previous research projects carried out by the author indicate that the connectionist network with a single hidden layer provides substantially improved modelling capacity over traditionally employed methods such as logit, being able to extract nonlinear patterns from the data (Greene, 2011). It is important to keep in mind that this relatively simple network architecture is nevertheless extremely powerful and capable to solve complex tasks in efficient and effective manner as shown above.

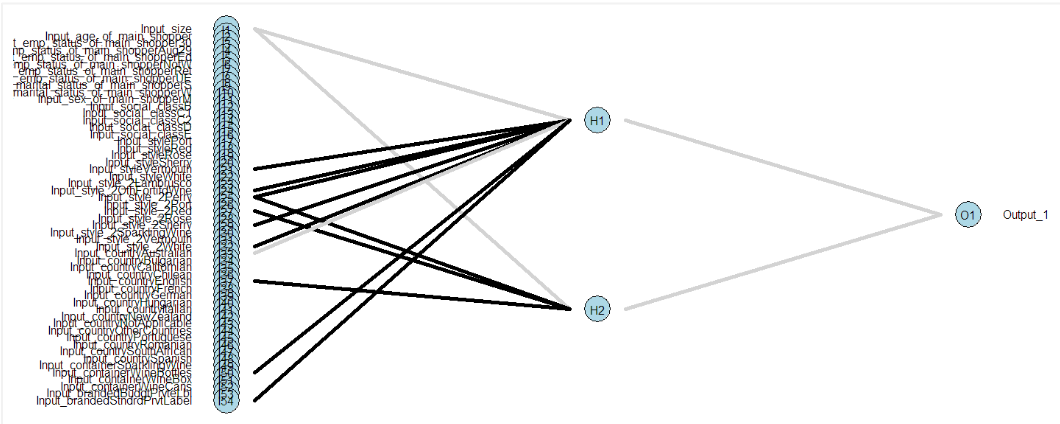

Figure 28. Connectionist network 54-2-1 architecture using consumer data with a single hidden layer and 2 neurons, 10 000 iterations, pruning with 100 retrain cycles.

There are a number of ways possible however to increase the model complexity while at the same time maintaining the underlying agenda to develop the network architecture open to interpretation. One way is to increase the number of iterations and retraining cycles, allowing the network to learn everything there is to learn within the given network architecture with only 2 hidden node. Network shown in Figure 28 is identical to the one in Figure 27 with the exception of increasing the number of iterations from 1000 to 10 000 – it takes only 42.07 sec to run and as a result pruning algorithm removes all but 15 connections, making it even easier to study and use the network architecture to explain the relations within the data.

5.7 Connectionist model predictive capacity

The original comparison of a regression model with the neural network model in section 5.3 can now be revisited and supplemented with the assessment of predictive capacity that a neural network with hidden layers and pruning is able to offer. Previous research programme (Greene, 2011) focused on a comparative assessment of a traditional method of analysis represented by a logistic regression, which is widely common in the marketing and consumer behaviour literature, with a connectionist model in the form of a feed-forward neural network with a single hidden layer – as a result, neural network model shows superior predictive capacity than a logistic regression model. For that reason, the emphasis here would be to assess predictive capacity of various connectionist architectures only, focusing on pruning capacity. Logistic regression is only able to show level of performance comparable to the simplest network architecture with no hidden layers as shown in section 5.3, and therefore is not considered here.

Using the same dataset as in the analyses carried out in section 5.6, a number of neural network architectures are assessed in terms of predictive capacity and model fitness.

Connectionist network 54-4-4-4-1 architecture is assessed, examining unpruned and pruned variants. Limiting to 1000 iterations, network models without pruning take as little as 54 sec to run, and produces 252 connections between 67computational units. Once pruning algorithm is introduced, it models take substantially more time to run – one of the lowest at 660 sec, with most running into 1000+ sec. As a result however, networks with as few as only 29 connections are developed, making it considerably easier to examine and interpret the network architecture and connections. While examining the RMSE figures for networks with and without pruning, the differences are comparable: networks without pruning show RMSE figures as low as 0.79, while models with pruning show RMSE figures that are able to reach levels as low as 0.82, which is very comparable considering only 1000 iterations were used to develop the connectionist models.

When connectionist networks with 54-6-4-2-1 architecture are examined in the similar manner, the observed results are similar to the 54-4-4-4-1 networks. Networks without pruning and 1000 iterations take at least 58 sec to run, and produce 358 connections between 67 units. Network design with architecture 54-6-4-2-1 allow more connections between the input layer and first hidden layer which here contains 6 rather than 4 units in 54-4-4-4-1 architecture, and therefore provides higher total number of connections. Once pruning is introduced, it takes 1676 sec to produce a network with substantially simplified architecture (certain model runs take even less time, but fail to reduce the network architecture complexity to the similar level, and thus are omitted here).

As a result, it is possible to reduce network architecture to as few as 47 connections amongst 62 computational units, which again provides a substantially simplified network for subsequent examination and interpretation. RMSE figures for unpruned 54-6-4-2-1 networks could be as low as 0.83, whereas models with pruning show RMSE figures that can reach 0.87 – again, quite comparable performance, especially considering that only 1000 iterations were used to run the models.

When looking at bottleneck networks with 54-2-4-6-1 architecture, models that do not use pruning produce 46 sec to run with 1000 iterations, and produce 146 connections between 67 units. Number of connections is substantially lower than in 54-6-4-2-1 and 54-4-4-4-1 network architectures – again, it is the reduced number between the input layer that contains the majority of neurons with our data, and hidden layer that now contains only 2 units that provides lower number of total connections as a result. Introducing pruning increases model run time to 311 sec, and reduces the number of connections to as few as only 9 connections, pruning out most hidden layers as a result. Occasionally the model would take substantially less time to run, but as a result it would fail to reduce the architecture complexity to the full potential – generally speaking with models that use only 1000 iterations, the longer it takes to run, the simpler the network architecture would be as a result, suggesting that pruning requires certain computational effort to be carried out properly. As a result, networks without pruning show RMSE starting at 0.90, whereas networks with pruning show RMSE figures as low as 0.95.

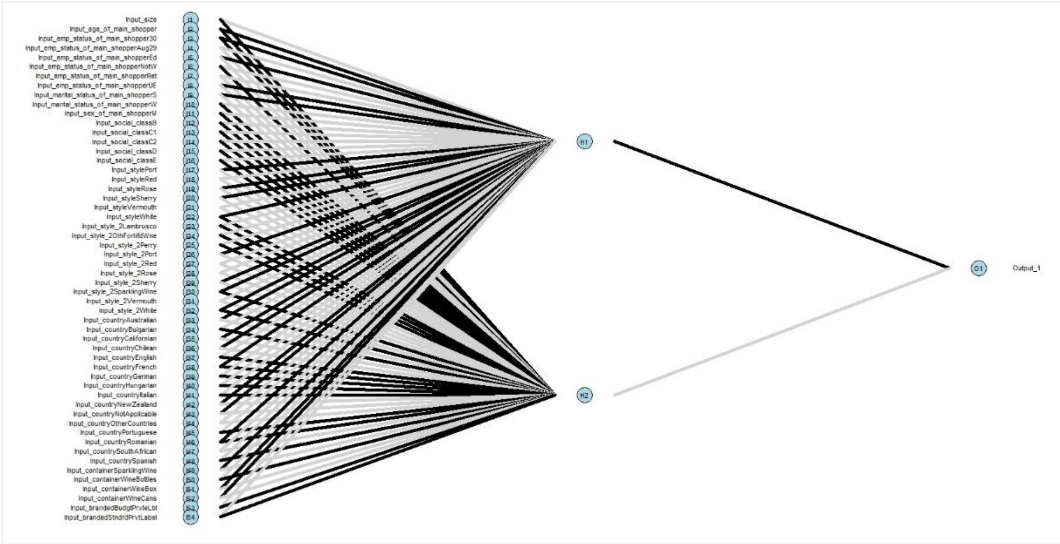

Figure 29. Connectionist network 54-2-1 architecture using consumer data with a single hidden layer and 2 neurons, 10 000 iterations, no pruning.

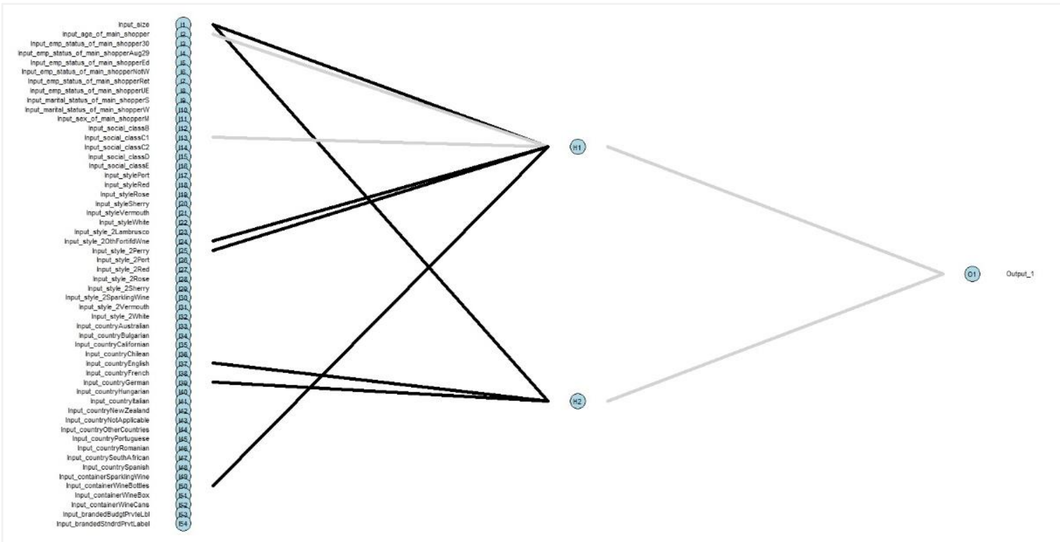

Figure 30. Connectionist network 54-2-1 architecture using consumer data with a single hidden layer and 2 neurons, 10 000 iterations, pruning with 100 retrain cycles.

Considering a substantially simplified network architecture as suggested in the previous section with 54-2-1 design, it takes as little as 38 sec for network without pruning to go through 1000 iterations, producing 110 connections between 57 units. Pruning 54-2-1 network takes 175 sec to run and can eliminate majority of connections, leaving fewer than 10 connections. RMSE figure of 0.85 can be observed with models when pruning was not used, and when pruning is carried out RMSE figures of 0.88 are possible.

When number of iterations is increased to 10000, the results over multiple model runs are more consistent compared to 1000 iteration models, but do not seem to show better RMSE figures. Increasing number of iterations to 100000 does improve RMSE figures. It also becomes apparent that the difference in time it takes to run models without and with pruning is reduced with more iterations, suggesting that pruning more complex models with higher number of iterations takes only marginally more time, which is a very positive feature in terms of scaling model size.

It is then should be apparent that pruning is able reduce the number of connections and network architecture substantially, while maintaining comparable level of predictive capacity. The ability to expose the core architecture of the data after all possible interactions were explored during model training would greatly simplify the task of exposing relations within the data that can be used not only for interpretation and explanation (as can be seen by comparing Figure 29 and Figure 30), but also for more transparent predictive modelling.

5.8 BPM connectionist model

So this brings us to the main question which is as follows: would it be possible to train a connectionist network that provides a sufficient predictive capacity, and then use the connection weights and hidden nodes as a distributed representation of informational and utilitarian reinforcement to develop the explanatory model sufficient for interpretation of the decision-making process as proposed by the theoretical framework of BPM? In addition, would pruning the network architecture reduce the model complexity and as a result provide a clearer explanatory framework while offering a comparable level of predictive capacity? In the following section, this is explored in detail.

5.9 Summary

In this chapter, results of the statistical testing methods employed as part of this research project were presented. First, results of preliminary data manipulations and exploratory analysis were described. Then regression and connectionist models were compared to establish a connection between the two methods of analysis on a computational level. This was followed by a discussion of the results from connectionist models of varying complexity. Finally, the capacity of pruning algorithms to optimise the network architecture and expose the predictive core were assessed in its attempt to provide a plausible account of representing informational and utilitarian reinforcement within the hidden layers as part of emergent distributed network architecture.

In the following chapter, results are discussed and interpreted.

6. Interpretation of results

In this chapter, the obtained results as described above are discussed and interpreted within the wider context of research questions posed here, and the field of consumer behaviour in general.

6.1 Informational and Utilitarian Reinforcement

For the sake of clarity before we proceed with the discussion of results, and to restate how the concept of utilitarian and informational reinforcement is operationalised and defined here following the connectionist frame of inquiry, it would be useful to summarise how informational and utilitarian reinforcement were examined in conjunction with connectionist modelling as part of the research programme to date. Prior to this research project, the theoretical approach was substantially different from the course identified here, as informational and utilitarian reinforcement were previously introduced into the neural network model in the form of additional input variables (Greene, 2011), and even though results demonstrated that doing so significantly improved model performance and corroborated previous findings (Foxall, Yan, Oliveira-Castro, & Wells, 2013; Yan, Foxall, & Doyle, 2012a, 2012b), subsequently it was identified and proposed that an entirely different modelling approach was required to explore the concept of utilitarian and informational reinforcement using connectionist networks (Greene, 2011). It was speculated that one reason for such improvement could be due to the process of assigning the informational and utilitarian reinforcement values to each brand available for selection, as it effectively allowed to capture in the quantitative manner some of the information on consumer decision-making setting and learning history which possibly be lost in otherwise the process of transforming the brand-level data for statistical analyses. This time however as part of the present research project, the informational and utilitarian reinforcement are not added on the input variable level as before, but rather are examined as a part of the artificial neural network learning process on the level of connectionist weights, and it is argued that the informational and utilitarian reinforcement are formed during the model learning process following the principle of distributed representation, and thus are the emergent entities within the hidden layers of the connectionist model.

6.2 Results discussion

Consumer behaviour modelling employing the NNs is able to offer as a result a substantial amount of information open for interpretation and further analyses. In the following sections, these essential aspects will be discussed in detail.

6.2.1 Exploratory data examination

It is important to note a few points regarding the data employed here: the synthetic datasets developed to test the modelling capacity, and the actual consumer data that contains transactional purchasing information, household descriptor database, and product attribute database.

Considering the vast amount of data available in the *wine* subset of the consumer dataset that was obtained from the Kantar World Panel, it is quite reasonable to expect that it would have been sufficiently robust for the purposes of this research to develop and test consumer models with sufficient level of predictive and explanatory capacity. Indeed, this was the case in previous research (Greene, 2011) where only consumer data was employed for comparative purposes to assess the adequacy of traditional and connectionist approaches to model and predict consumer situation and consumer decision-making faculties. This time however, as nonlinear connectionist models of higher complexity with multiple hidden layers developed here require new software solutions to be developed in parallel to the research process, these very algorithms themselves needed to be assessed at the outset, before they can be employed to examine and model consumer data. Thus, additional synthetic datasets were devised, which would allow the relations in the synthesised data to be defined by design. Relations in the consumer data, on the other hand, are not defined and thusly are not suitable to be used to assess the adequacy and performance of algorithms and software solutions – it is in fact the overarching purpose of this very research project, to identify and extract the patterns in the consumer data, which could then be useful to explain consumer behaviour and purchasing decision-making process.

6.2.2 Results of regression and NNs comparative analysis

It is of course a widespread misconception that NNs models do not provide sufficient explanatory capacity as by design they lack the computational and processing transparency, and thus are not able to offer to researchers a glimpse into the model development process. This may complicate the capacity of researcher to interpret and explain the results thus obtained, but it is commonly assumed that researchers may accept this limitation because the model is able to offer predictive superiority as a sufficient trade-off. It is however not the model inherent design that makes it difficult to interpret the results, but rather the intricate nature of the relations within the data that the NNs model aims to unravel that human researchers may struggle to comprehend and explain in simple terms. In fact, this is the very reason why researchers employ nonlinear parallel modelling in the first place – to decompose complex phenomena and reduce the dimensionality of the problem and make it easier to comprehend and interpret.

Moreover, it is argued here that connectionist models are not only adequate to explain the relations between variables within the data as compared to traditional regression models, but are superior not only in predictive but also in explanatory capacity when the research problem and data complexity is high, thus refuting the *back box* misconception commonly attributed to connectionist modelling. One major reason for this argument is of course the inherent capability of connectionist models to provide a comprehensive account of nonlinear interaction between the variables in the data, something where traditional methods could have serious capacity limitation – in particular an issue with scalability when large datasets with many variables are involved, as number of interactions to consider increases exponentially with the number of variables.

Connectionist networks, on the other hand, are able to deal with all interactions within the data as part of a learning process, while the network architecture is developed to represent the patters within the data rather than being pre-specified by a researcher as require with traditional methods of analysis.

To address these points, traditionally employed in marketing and consumer literature regression analysis is compared with the NNs model of the simplest network architecture. The problem type is transformed to a dichotomous for this particular task only – to align the distinctly different modelling approaches to an even level, as it is not the actual model performance that is being compared at this point, but rather the ability of the connectionist model to offer a level of performance equal to that of its corresponding traditional model. As a result of a series of comparative assessments that involve multiple random splits and replications of the procedure to increase the validity of results, it is clear that the connectionist model with the simplest network architecture that includes no hidden layers is able to perform identically to a regression model – the method of choice in traditional marketing and consumer literature and fields of study.

Connectionist models with simplified architecture with no hidden layers developed and examined in the first stage of research project here were able to show performance levels equal to performance of logit models: NNs models that comprised solely of the input and output nodes, and incorporated no hidden nodes, were able to offer connection weight values identical the coefficients in logit models. No hidden layers of course means the simplified straightforward NNs models were not able to account for any nonlinear relations within the data

– it can be said then that connectionist models display not only the predictive performance levels equal to that of logit model counterpart, but also provide a equivalent level of explanatory capacity on variable contribution dimension. Thus, a clear link is established between the connectionist architecture of NNs and the traditionally employed logit modelling as a method of analysis while working with consumer data, and offers a solid starting point from which the two methods of analysis can be shown to contrast substantively both in underlying architecture and design, and levels of performance. The following paragraphs outline a number of important points to support this.

First imperative point to discuss is the required predetermined structure of a logit model, and the lack of thereof by design in connectionist models. The simple models developed at this stage could of course be advanced and developed further as required – both logit and NNs. Common way to improve a logit model is to examine the variable contributions and pick the optimal set that includes the variables that offer best predictive capacity, while at the same time introducing the interactive variables to account for possible nonlinear relations within the dataset. This task however is rather manual and tedious, and requires not only computational resources, but also human resources as it may take researcher a considerable effort to continuously test and assess alternative models to select the optimal predictive set. Additionally, large datasets with numerous variables to consider pose an even larger issue as number of possible interaction combinations increases exponentially. As noted elsewhere (Bishop, 1995), the fact that traditionally employed methods of analysis require a predetermined model structure is an inherent weakness when compared with the connectionist models. Whereas connectionist models, on the other hand, are able to determine its own structure and network architecture following a statistical method (as much as this term can be applied to artificial entities) during the modelling process while extracting the patterns from the dataset. Therefore, to increase the capacity of the NNs model and advance its complexity and capacity, it is only a matter of increasing the number of computational nodes and (hidden) layers – the optimal architecture is then determined during the network learning process and there is no need for researcher to specify or define it beforehand. The increased network complexity by introducing hidden layers introduces the network capacity to account for nonlinearity, which considerably improves the performance of the NNs model with complex consumer data.

Second, it is important to touch upon the theoretical implications that the predetermined model architecture may postulate. As discussed above, variable selection process to form the model structure is not only tedious and taxing, but also could be studied as a research question in its own right (for example see Greenland, 1989). Underlying theoretical or philosophical framework could dictate the models structure as well. It is however a common practise with traditional methods of analysis such as logit to develop the model structure and carry out the variable selection process often relying on predictive capacity alone to be used for all subsequent analyses – generally a goodness of fit criteria such as *R-squared*, or another purely statistical measure. Depending on the underlying theoretical framework, this may bring either positive or negative consequences.

On the one hand, the predetermined model structure may be used to control the exclusion or inclusion of variables that may be of particular interest during the operationalisation or research questions. The pre-programmed model structure could also make the process of understanding and explaining the phenomenon somewhat more straightforward. On the other hand, it is safe to assume that the predetermined model structure would inevitable have an effect on the interpretation of the results obtained this way. Thus, the positivistic aim for a researcher to be largely removed from the subject matter could be said to be compromised – in extreme cases rendering the results and meaning derived from such results as nothing but a statistical artefact (Harris & Hahn, 2011). The nature of predetermined model structure may play even larger role depending on the type of research – for example, research questions that aim to provide a descriptive account of a process of the phenomenon would be greatly influenced if the model structure is in fact determined or even selected by the researcher, potentially compromising the overall objective, rather that developed to reflect the patterns in data. For such tasks, it should be more appropriate to employ such method of analysis where the model structure and underlying architecture is a product of analysis rather than an initial requirement – such as NNs models.

Third, the results described here address and carry a serious challenge to the *black box model* claim. In marketing literature discussions and in the industry, it is frequent to encounter the arguments against connectionist models that revolve around the difficulties to explain and interpret the modelling process: the model manipulates the input data according to certain rules and provides an output, but makes it difficult to examine and interpret the modelling process with simple terms (Gevrey, Dimopoulos, & Lek, 2003; Olden & Jackson, 2002; Olden, Joy, & Death, 2004). As discussed above, the empirical evidence offered here suggest otherwise and argues this claim to be untrue: if simplifies input-output connectionist model is capable of generating results identical to those of a logit model, it should be reasonable to infer that the explanatory power that simplified connectionist model carries is comparable in the very least with the explanatory capacity attributable to a logit model. The fact that the connection weights offered by the NNs model are identical to the coefficient values of the logit model supports this argument, and explanatory capacity of NNs models could be considered to be on the same level as regression model is able to offer. Complex connectionist models follow the same theoretical assumptions and statistical rules as simplified model described above – it is these NNs models of higher complexity that are often referred to as black box models, as complex patterns extracted from the data become less obvious to researchers, and therefore less straightforward to explain and interpret in simple terms. Nevertheless, this does not point to the flaw of the model but rather to the boundaries of human understanding, and the limited ability to consider complex multi-dimensionality, and thus the *black box* analogy is hardly suitable for connectionist models.

In summary, simpler problems do not require complex algorithms, and are easier to explain and interpret. Simple algorithms are inadequate when applied to complex phenomena however, which require complex algorithms and models – but complex phenomena are difficult to explain and interpret no matter what algorithms and models are employed, as it is the complex nature of the problem that makes it difficult to comprehend, not the method employed. This is a fundamental notion here, as all consecutive modelling development follows this very same principle by adding the necessary complexity to the connectionist network architecture as dictated by the increasing complexity of the research questions, and the need to examine relations in the data and phenomena of higher informational density.

6.2.3 Exploratory NNs modelling results and discussion

Now that it is established that simplified NNs models are able to perform on par with logit models and provide analogous level of explanatory power, the connectionist models are developed further. Introducing additional hidden layers between the model input and output provides the ability for the connectionist model to account for nonlinear relations in the data. It should be safe to assume that transactional data on consumer purchasing decision in a particular marketing setting is complex, and therefore it is expected to contain a substantial extent of nonlinear relations – models able to account for such relations should provide considerable advantage over linear models in both predictive and explanatory capacity. Low *adjusted R* described above may indicate relatively weak logit model performance due to a number of factors. There is of course always a possibility that the particular dataset does not in fact carry any predictive capacity and is therefore not particularly useful in answering specific research questions irrespective of the statistical methods employed to analyse it. In fact, even though connectionist models do not require a predefined model structure and architecture, researchers may be limited to what can be included nevertheless as the data collected may have followed a particular philosophical position or framework – whether explicitly stated or not. Thus, the data collected in such a way would inherently only include the variables according to this philosophical position or framework, limiting the ability of connectionist model to extract patters from the data from the onset. Here however secondary data is used, and the number of variables included is very large to say the least. Another possibility for the relatively weak performance of logit model could be that the relations between dependent and independent variables that are being described are in fact not linear. If so, appropriate method of analysis should be able to improve upon the results obtained with logit models by being able to account for nonlinear relations within the consumer dataset.

Previous research carried out by this author examined the effect of network complexity provided encouraging results – the focus at the time was on the number of hidden neurons, and was limited to a single hidden layer (Greene, 2011). The dataset employed in the previous research programme was sufficiently large (around 75 000 cases for a single product category) to make sure connectionist network with many hidden neurons are not able to learn the entire dataset to attain a nearly flawless prediction power. The effect of model size on model performance was examined with networks with size starting from a single hidden neuron and progressively increasing to a 100 hidden neurons total (and even a few select models that used as many as 200 neurons within a single hidden layer) and compared with the performance of a logit model (which of course could not be developed further beyond the initial input variable selection). The results showed gradual performance improvement as the connectionist network size increased and more neurons were employed in the network architecture, providing better means for the model to identify the nonlinear patterns in the data. This suggested that connectionist models were suitable to study complex datasets and consumer behaviour data in particular.

It should be expected for the model performance to flatten out at some point, where increasing model size would not provide substantial improvement in the model capacity – it was not observed with the dataset employed in the previous research project (Greene, 2011), and as the dataset employed here is considerably larger and more complex yet, it should be safe to assume this limitation would not be an issue here either. It was then further discussed that with sufficiently large datasets it would be beneficial to extend the scope of research project to explore performance with multiple hidden layers – something that may easily be a substantial research project in itself. This was of course for the most part an exercise to assess the relation of the size of connectionist network and predictive capacity; whereas here it is the explanatory capacity which the network architecture may be able to provide is the focus of research questions, and thus the notion of assessing the network architecture performance with multiple hidden layers will be address only partially and from an explanatory modelling point of view.

6.2.3.1 RSNNS and NeuralNetTools

It is noteworthy to remark upon the importance of collaboration within the scientific community in the field of social science that not only builds upon the previous work and research findings, but also goes into the wide-ranging efforts of tools development. As should be obvious at this point, the necessary tools to carry out the complex statistical and computational methods were absolutely crucial to the investigation of the research questions posed here, and it is the collaborative nature of researchers from various fields that made it possible. Moreover, as a result now, the advanced tools and methods are available to any and all other researchers out there through a free platform and statistical environment.

6.3 Advanced connectionist models results and discussion

In the following sections, the investigative analyses that focused on comparing the connectionist network architectures of various sizes from both predictive and explanatory point of view are discussed in detail.

6.3.1 Connectionist network predictive capacity

Previous work showed connectionist models to be vastly superior to traditionally employed forecasting and predictive analysis methods such as logistic regression (Greene, 2011). A number of analytical directions were discussed as part of previous research programme that are now able to contribute to the present research which builds upon those findings and focuses on the identified and proposed areas of further development. Consumer data was employed at that time, thus the findings should be applicable for the purposes of present discussion – even so, additional analyses were carried out here as well nevertheless as described in the previous chapter, addressing some of the proposed in the previous research areas of potential development (Greene, 2011).

As identified and proposed in the earlier research programme, the nature of the task revolves mainly around the concept of optimal architecture selection, and could be recognised as twofold: on the one hand, the methodological procedure to explore and assess the extent of potential network architectures is required; and on the other hand, a decision mechanism which could be employed in architecture selection process. Considering the model generalisation capacity and the primary function of the connectionist model, one of the common approaches to determine the optimal network architecture is a deliberate exploratory analysis where a limited number of suitable architectures are explored and assessed, such as the analyses carried out in previous research that examined and compared the predictive capacity of the connectionist network depending on the number of hidden neurons within a single hidden layer (Greene, 2011), which have been extended here as described in the particular sections above that examine the effect of the size and structure of the connectionist network with multiple hidden layers on the overall network predictive and explanatory performance. This approach however requires a considerable computational effort to explore a limited group of networks, which could present a number of limitations as a result. Moreover, using a broader group inclusion criteria as employed here to study the connectionist models with varying architecture structures and multiple hidden layers increases the feasibility of reaching the computational limit (Bishop, 1995). In fact, many models described here have been considerably restricted by limiting the network and sample size due to computational limitations, otherwise resulting in consistent crashes of some of the algorithms employed here. Additionally, another obvious drawback of this approach is the requirement to train a large number of networks with varying architectures to compare and select the optimal architecture that satisfies primary and secondary requirements, be that predictive or explanatory capacity.

A considerably better approach that would be able to account for some of these limitations at least to some extent is to remove computational neurons and synaptic connections from a large complex network architecture in a systematic manner, exposing the core network architecture while minimising the network size – these methods are commonly referred to as *pruning* algorithms (Bishop, 1995). Pruning method employed here is the *Optimal Brain Surgeon* as described in previous section and elsewhere (Hassibi & Stork, 1993; Hassibi, Stork, & Wolff, 1993) in greater detail.

Alternatively, it is possible to go in a reverse order and start with a relatively simple connectionist network only to sequentially expand its architecture by gradually adding more computational neurons that would form new connections and hidden layers – a group of method commonly referred to as *growing algorithms* (Bishop, 1995). Potentially this could be combined with the pruning method employed here: growing the initial complex network architecture and subsequently pruning it to expose the core structure that carries the best predictive and explanatory function.

Another approach involves aggregating a number of networks to function together as a single entity – this method is commonly referred to as a *network committee*. It employs a divide and conquer strategy in which the response of multiple constituent networks is combined to provide a superior single expert response (Bishop, 1995) – particularly suitable to study phenomena that allow themselves to be decomposed and subdivided into smaller isolated tasks to be examined separately.

Number of computational units and the synaptic topology are able to exert a considerable influence over the network performance, and these connectionist network architecture optimisation methods would be expected to improve the modelling process significantly and provide means to develop a network architecture with superior predictive and explanatory abilities while using the optimal network structure and size.

6.3.2 Variable contribution analysis and explanatory models

From the discussion above it should be clear that connectionist models offer substantive predictive capacity and pattern recognition function. This research project however is primarily concerned with the explanatory dimension connectionist networks are able to offer.

Previous research programme offered a rather limited review as the explanatory dimensions of connectionist models was largely out of scope of the project (Greene, 2011), whereas here it is the primary focus. Variable contribution discussion would make sense here – on the one hand, to address further the argument that connectionist models are *black box* models, and on the other hand, to explain the logic in undertaking the sequence of analytical account that is to follow in the next sections. The black box argument is of course completely unwarranted as already discussed to some extent earlier, and variable contribution analysis is another good approach often employed to further refute these unfounded claims. One such framework was proposed in the field of ecological research (Gevrey et al., 2003; Olden & Jackson, 2002; Olden et al., 2004) and this author previously proposed to evaluate its usefulness with consumer behaviour data. Essentially, it can be contended that connectionist models are able to provide a comparable level of explanatory capacity with the traditionally employed methods such as regression analysis. Even more so, it can be argued that it is in fact the connectionist methods that are able to provide a robust account of behaviour when it comes to the explanatory dimension – on the contrary, it is the traditionally employed methods such as regression analysis that are not able to provide an adequate explanatory of behaviour. In regressions for example, normally the partial regression coefficients are assessed to interpret variable contribution – this is only true for statistically significant variables however, and even then very little is provided by the model besides the coefficient value and the sign for the independent variables that signify the direction of the relation, while no further information could be extracted. Whereas the situation is quite the opposite with connectionist models – variable contribution analysis algorithms could be employed to provide the explanatory account and determine relative variable contribution and contribution profile of the input factors. In addition to pruning algorithms which are discussed in great detail in the following sections that improve both predictive and explanatory capacity of the connectionist networks as part of learning process, there are also variable contribution algorithms for connectionist models that focus on explanation and interpretation only and aim to estimate the relative contribution of independent variables and determine what relative contribution each input variable is able to provide in relations to the output of the model.

Gevrey, Dimopoulos, and Lek (2003) examined 7 methods that potentially could be useful to perform variable contribution analysis with consumer behaviour data: (1) the *PaD* method calculates the partial output derivatives according to the input variables (Dimopoulos, Chronopoulos, Chronopoulou-Sereli, & Lek, 1999); (2) the *Weights* method uses the connection weights (Garson, 1991); (3) the *Perturb* method performs input variable perturbation (Scardi & Harding, 1999); (4) the *Profile* method is a successive variation of one input variable while others are kept at a fixed value (Sovan Lek, Belaud, Baran, Dimopoulos, & Delacoste, 1996; S Lek, Belaud, Dimopoulos, Lauga, & Moreau, 1995; Sovan Lek, Delacoste, et al., 1996); (5) the *classical stepwise* method observes the change in the mean square error value by sequentially adding or removing input neurons (Maier & Dandy, 1996); (6) *Improved stepwise - method a* is the same as (5), but the input eliminations occur while the network is trained (Gevrey et al., 2003); and (7) *Improved stepwise – method b* evaluates the change in the mean square error by sequentially setting input neurons to the mean value (please see Gevrey et al., 2003). They used a multi-layer feed-forward network architecture to compare the variable contribution analysis methods – as a result, all methods had the ability to order the variables by importance of their contribution to the output, but the *PaD* and *Profile* methods were able to also order of contribution and mode of action. The *PaD* method employs partial derivatives and real dataset variables values and was identified as the most robust and coherent computationally, followed by the *Profile* method where a representational matrix of the data is constructed.

It should be safe to assume that ecological systems are complex enough to provide data with complex relations, so the promising results that the techniques surveyed by Gevrey and colleagues show should be applicable to consumer behaviour data as well. Many additional factors should be taken into account of course – the more obvious one is whether these variable contribution analysis methods would be able to perform on a comparable level with much larger sample sizes that contain a multitude of input variables. Relatively small sample size and the 10-5-1 network structure used by Gevrey et al. is largely different from the datasets used here, but nevertheless there are some important learnings that can be extrapolated from the study as far as the further developments, and it potentially this could be an important area to consider in future work with employing the variable contribution analysis algorithms to a pruned network architecture to assess the effectiveness and efficiency of the algorithms, and whether they are able to contribute to explanation of behaviour after pruning connectionist network. It is of some concern however that the real ecological data was used in the comparative study to assess the variable contribution algorithms – similar concerns expressed elsewhere (Olden et al., 2004) contend that true relations and variable order are not known in the empirical dataset employed by Gevrey et al. and therefore there is no straightforward way to assess the accuracy that each method is able to offer. Some of the other methodological issues were identified with the research design by Olden and colleagues (2004) which prompted Gevrey et al. to redesign the original study and use a synthetic simulated data instead to re-assess the variable contribution analysis methods, adding an additional *connection weights* algorithm (Olden & Jackson, 2002) at the same time. This time around, *connection weights* algorithm was identifies as optimal as a result, showing the best level of performance with simulated data.

The variable contribution analysis could be a great area for further research to explore elsewhere, but it is out of scope of present research project. The concept of using the simulated data however is not, and it makes all the sense to for the research design perspective to incorporate the simulated data for the initial assessment of the pruning efficiency and effectiveness in attempt to simplify the network architecture for consecutive interpretative and explanatory account of consumer behaviour and decision-making process. In the following sections, it is the synthetic simulated data that is reviewed and evaluated first, which is then followed by the discussions of the models that employ real consumer behaviour data.

6.4 Interpreting connectionist model output parameters and architecture

A number of ways may provide an insight into what happens inside the NNs model and help interpret the result. Some of the most commonly used methods assess how the number of hidden layers and nodes affects the predictive and explanatory capacity of the model. A number of algorithms have been devised to make use of the weight values from NNs model output. Model architecture pruning techniques have also been shown to have a positive outcome in developing models with improved out of sample testing faculties. In the following sections, these methods will be briefly discussed and supported by the empirical research.

6.4.1 Number of hidden layers and nodes

Generally, a larger model would have better resources at its disposal to analyse extensive datasets and show better predictive capacity, but would also have considerable limitations as discussed in the following paragraphs.

6.4.1.1 Model structure optimization

Once the models are developed it is imperative to have a look into the optimal model structure. It is indeed true that the larger models would offer higher predictive capacity and increase in the model fit, but at the same time, larger models need to be penalized according to the Occam's razor principle. One method to evaluate the model performance and select the optimal structure is described by Huang, Chen, Hsu, Chen, and Wu (2004). Before carrying out the analyses that would employ the neural network method for modelling, the authors optimised the backpropagation models for multiple markets by identifying the optimal input variable sets that included financial variables following an approach that would resemble a process similar to that of a step-wise regression model: once a simple initial model was constructed to represent the financial markets, financial input variables were removed one at a time and replaced with a remaining variable in attempt to examine the effect it would have on the overall predictive capacity of the model. This process was repeated numerous times until improvement could no longer be observed – the final neural network architecture obtained in such a way was said to be optimal for each of the financial markets. When these models were tested with 10-fold cross-validation method, the prediction accuracy that these 2 models were able to offer was estimated to be optimal as well, and these 2 fine-tuned models were then used for all consecutive interpretative analyses (Huang et al., 2004). This method eliminates the independent variables that carry the least predictive and explanatory capacity and therefore can be excluded altogether from consideration in the model or replaced with other potential input variables to improve the overall predictive capacity. Thus, the model structure is simplified and therefore is more preferable – it is expected to show higher AIC values as well, as method described above penalizes model size to keep the connectionist model architecture as simple as possible while at the same time striving to maximise the overall model performance.

Even though apparently effective, it is difficult to estimate how effective it really is, as it was tested using the empirical data where the relations between the variables are not known, therefore making it practically impossible to assess the method efficiency and effectiveness at simplifying the network architecture while at the same time maintaining comfortable level of predictive and, more importantly for the purposes of research project here, explanatory capacity. Moreover, the process seems to be somewhat manual still, and potentially may require substantial effort to sequentially and systematically test yet more and more variables – the stopping mechanisms are not clearly defined either. It may seem the researchers may be faced with a similar issue as with input variable selection for the regression analysis, where in attempt to achieve high level of validity and test a large number of variable combinations and interactions, the truly robust process may prove to be excruciatingly taxing on both researcher time and other resource allocation. Moreover, this approach of course cannot be scalable in any reasonable manner – as dataset get larger, the amount of effort and resources required increases exponentially. In this sense, the elegant pruning algorithms could be a considerably more favourable solution to optimise the connectionist network architecture in attempt to streamline the interpretative and explanatory functions of the connectionist model.

6.4.2 Interpreting model weights

A number of methods have been shown to be useful in interpreting the weight values of connectionist models.

Variable contribution analysis methods have been examined and compared by Gevrey, Dimopoulos and Lek (2003). One of the seven methods they surveyed included a computation that used connection weights to provide explanatory dimension to a NNs model using ecological data. First proposed by Garson (1991) and later further investigated by Goh (1995), the procedure is set to determine the relative importance of the inputs by partitioning the connection weights. Essentially, hidden-output connection weight of hidden neurons is partitioned into components associated with the input neurons (for further details see Appendix A of the Gevrey et al., 2003). Authors concluded that method that uses connection weights was able to provide a good classification of input parameters even though it was found to lack stability.

One of the concerns conveyed regarding the otherwise extensive investigation of different methods was that the dataset originally employed in 2003 study (Gevrey et al.) was empirical, and therefore did not allow to ascertain the factual precision and accuracy of each method as the true relations between the variables are not known (Olden et al., 2004). Instead, the artificial dataset was created using the Monte Carlo simulation and employed to assess true accuracy of each method using the dataset with defined and therefore knows relations. Results show that *weights* method that uses input-hidden and hidden-output connection weights displayed consistently best results out of all methods assessed, contrary to Gevrey et al. original findings (2003). Additionally, the *weights* method was able to accurately identify the predictive importance ranking, whereas other methods were only able to identify the first few if any at all (Olden et al., 2004).

Olden and Jackson (2002) also used ecological data to demonstrate the predictive and explanatory power of NNs. A number of methods surveyed, including *Neural Interpretation Diagram*, Garson's algorithm, and sensitivity analysis, aid in understanding the mechanics of NNs, and improve the explanatory power of the models. Interpretation of statistical models is imperative for acquiring knowledge about the causal relationships behind the phenomena studied. They also propose a randomization approach to statistically evaluate the importance of connection weights and the contribution of input variables in the neural network – method discussed in further details in the sections above.

Nord and Jacobsson (1998) have also addressed the issue of explaining and interpreting NNs structure and developed algorithms for variable contribution analysis. The study compared the proposed novel algorithmic approach for NNs model interpretation with the analogous variable contribution method of partial least squares regression. Sensitivity analysis is also performed through setting each input to zero in a sequential manner. Linear regression coefficients for each of the input variables have also been generated for the purposes of examining the variable contribution direction. The results of the two approaches are then reviewed and compared with the results of the partial least squares regression. What the study is able to reveal is that in the linear dataset both the partial least squares regression and NNs models show similar performance in the variable contribution task, whereas with the nonlinear dataset the differences in performance is apparent (Nord & Jacobsson, 1998).

The recently increasing interest in variable contribution in NNs models is understandable, as this information could be useful to develop the optimal NNs model structure or to enhance model explanatory capacity. The methods commonly used examine the connection weights of the NNs model, which are used to interpret the model performance – analysis of the first order derivatives of NNs model with respect to input units, hidden units, and weights. For a more extensive discussion on measures of relative importance and relative strength of inputs please see Garson (1991). It is also possible to use the connection weights in attempt to extract the symbolic rules to interpret the models. Huang, Chen, Hsu, Chen and Wu (2004) use Garson's contribution measures to assess the relative importance of the inputs in NNs three-layer backpropagation model. Even though Garson's method emphasises connection weights between the hidden and the output layers and does not consider the direction of the influence, a comparative analysis revealed the method to improve understanding of the financial process being modelled. The contribution analysis employing Garson's method identified the input variables contribution to the output variables, which increased the understanding of financial input factors in the NNs model (Huang et al., 2004).

Lek, Belaud, Baran, Dimopoulos, and Delacoste (1996) examined model response to each of the variables. Functions derived by the NNs models during the learning stages are very complex and pose a serious problem for each variable contribution analysis. One of the ways to cope with such issues is to isolate a complex phenomenon and separate it into smaller less complex phenomena to be examined independently. Authors propose an experimental method to examine the model response to each of the variables by applying typical variations to a single separate variable while the other variables are held constant. Using the environmental data, all but one variables were sequentially set to their minimum, first quartile, median, third quartile, and maximum values providing a response. The operation is then repeated for each of the variables, performing it *n* times, where *n* is a total number of variables (Sovan Lek, Belaud, et al., 1996).

Relatively few studies are carried out with the aim of developing methods for variable contribution analysis in NNs models in particular – perhaps at least in part due to seeming complexity of the task at hand. Song, Kong, and Yu (1988) have developed a partial correlation index method that employs sequential removal of variables one at a time. Results obtained under standardized training conditions are used to estimate the relation between input and output variables. Nord & Jacobsson (1998) proposed alternative method based on the sequential zeroing of weights. Andersson, Aberg, and Jacobsson (2000) examined two methods to study variable contribution in NNs models: (1) a variable sensitivity analysis and (2) method of systematic variation of variables. Variable sensitivity analysis is based on setting the connection weights between the input and hidden layer to a zero in a sequential manner, whereas the systematic variation of variables method is based on keeping the other variables constant or manipulated simultaneously. In the course of the study, it is shown that there is a high similarity between the method proposed by the authors for the variable contribution analysis in NNS models and the nature of the processes used to develop the synthetic datasets used. Thus, it is shown that the NNs models are suitable not only for the function approximation in nonlinear datasets, but are also able to accurately reflect the characteristic qualities of the input variables. As a result, a transparency of highly interconnected NNs models could be demonstrated in response to the 'black box' argument as well. Presented method is then able to generate information about the variables that could be useful in examination and interpretation of variable contribution and relations. Nord and Jacobsson's method (1998) mentioned above is based on the saliency estimation principles (such as brain surgeon) as it estimates the consequence of weight deletion on prediction error. The difference with the method proposed by Andersson, Aberg and Jacobsson (2000) is in the way estimation is carried out (theoretical calculation in saliency estimation methods as opposed to experimentally derived values offered by Andersson et al., 2000), and builds upon the findings of Nord and Jacobsson (1998). In the course of analysis, a systematic variable contribution analysis is carried out on a highly interconnected network structure, including the signal separation exercise, employing a number of synthetic and empirical datasets to provide additional information on the methods considered, including the ability to graphically reveal the variable interdependencies. Previous research is considered there as well that is based on the principle of systematic variable variation and not the connection weights. Information obtained in such a way could constitute an analytical basis for a comprehensive variable contribution analysis and variable selection procedure survey (Andersson et al., 2000).

6.4.3 Pruning connectionist models

Model architecture plays an important role in a model adaptive performance. The type of a task closely related to the connectionist model pruning effort that attracted larger interest in the literature as discussed above is variable selection method, which is mainly concerned with the methodical improvement of the connectionist model architecture by systematic reduction of the input variables (Andersson et al., 2000). There are a number of notions that variable selection methods could consider: for example a method that examines the connectionist model weight values (Ametller, Garrido, Stimpfl-Abele, Talavera, & Yepes, 1996) employing the variance and saliency measures. Other approaches considered employing various other methods such as F-test, principal component analysis, decision tree methods, connectionist weight evaluation methods (Cibas, Soulié, Gallinari, & Raudys, 1996; Proriol, 1995), and *optimal brain damage* algorithm (LeCun, Denker, Solla, Howard, & Jackel, 1989). Despagne and Massart (1998) discussed variable selection methods, and among a number of the different approaches reviewed, which include a modified variant of Hinton diagram, saliency estimation method, and two other methods that provide a means to estimate the extent to which the variance of the predicted response corresponds with the variable contribution of each input. Similar to the method proposed by Nord and Jacobsson (1998), both methods revolve around the notion of cancelling variable contribution in the trained connectionist network by either zeroing input variables (Despagne & Massart, 1998) or connection weights (Nord & Jacobsson, 1998).

While exploring how environmental conditions have an effect on fish population to identify patterns that may be useful in future predictions, Olden and Jackson (2001) compared traditional statistical approaches with NNs models. In the NNs mode structure, the connection weights between neurons are the associative links that signify the relation between the input and output variables and therefore are the key to solving the problem. Connection weights signify the influence each input variable is able to exert on the output, and dictate the direction of the influence. Input variables with large connective weights carry higher signal transfer capacity and therefore affect the output variable to a greater extent. Excitatory effect (incoming signal increased with positive output effect) is represented by the positive connection weight and inhibitory effect (incoming signal reduced with negative output) is represented by the negative connection weight. In recent work, some research supports the notion that it is possible to use the connection weights to interpret the input variable contribution in the task of predicting the network output (Aoki & Komatsu, 1997; Chen & Ware, 1999; Özesmi & Özesmi, 1999). Others used the connection weights to quantify the variable contribution ranking (Garson, 1991), or employ sensitivity analysis to examine the input variable contribution range (Guégan, Lek, & Oberdorff, 1998; Sovan Lek, Belaud, et al., 1996; Mastrorillo, Lek, & Dauba, 1997; Mastrorillo, Lek, Dauba, & Belaud, 1997). Even if it is possible to assess the overall contribution of input variables employing these approaches, the interpretation of interactive relations within the data presents an increasingly difficult undertaking, as the interactions between the variables in the network require immediate examination. Even a small network would contain a large number of connections, making the interpretation increasingly difficult: 10 (input) – 5 (hidden) – 1 (output) network would have 50 connection weights to examine between the input and hidden layers. One way to manage this is through pruning where connections with small weights that do not exert significant influence over the network structure and output are removed (Bishop, 1995). Deciding which weights to remove and keep however is a task that requires substantial effort. Following the connectionist approach, Olden and Jackson (2001) were able to develop a randomization test to address this task, which aims to randomise the response variables to subsequently proceed with constructing a connectionist network using this randomised dataset, at which stage all connection weights between the input, hidden, and output connection nodes are recorded. This process is replicated 10,000 times to ensure the estimated probability values are stable to obtain a null distribution for the input, hidden, and output nodes, which are then compared to the observed values – this allows to calculate the significance levels which serve as the basis for the objective pruning test that allows elimination of the connection weights that exert a minimal influence of the network overall output and performance, and as a result helps identify those input variables that are able to provide the best predictive capacity contribution to the overall connectionist network performance. In the similar manner as was carried out as part of the present research project, Olden and Jackson (2001) considered varying levels of learning rate and parameters during the connectionist network training stages to maximise the probability of global convergence, and also considered varying numbers of training cycles to identify the optimal level as far as network training and performance balances with the resource allocation and training times. All input variables used to develop the connectionist networks were standardised in the preliminary data manipulation and exploratory analyses stages to avoid any possible occurrence of unnecessary variances between the input and output variables due to the differing variable scales. As a result, Olden and Jackson (2001) were able to provide a predictive and explanatory insight into nonlinear complex relations of ecological data (a task that poses a serious problem for traditional statistical approaches as species often exhibit nonlinear response to environmental conditions). In the course of detailed evaluation of NNs and traditional models it was shown that partitioning the predictive performance of the model into measures such as sensitivity (ability to predict the presence) and specificity (ability to predict the absence) allows for a more efficient way to assess the model strengths, weaknesses, and applicability. It is also shown that NNs are a useful approach for examining the interactive effects and factors. Both empirical and simulated datasets were used for comparative purposes, and show superior predictive performance of NNs models over traditional regression approaches (Olden & Jackson, 2001). Building upon the work described thus far, approach that Olden and Jackson (2002) propose in their following publication provides the facility to eliminate irrelevant connections between neurons whose weights do not significantly influence the network output (i.e. predicted response variable), thus facilitating the interpretation of individual and interacting contributions of the input variables in the network. The approach is able to identify variables that provide a significant contribution to network predictive capacity, which effectively constitutes a NNs variable selection method.

One aspect worth discussing however is the approach to identify the optimal number of neurons to use within the hidden layer of the network architecture: it was determined by Olden and Jackson (2001) following the empirical investigation where the performance of connectionist networks of varying sizes (ranging from 1 to 20 hidden neurons) were compared to identify and select the one with the network architecture that offers best predictive capacity for the overall connectionist model. This approach is similar to the research work carried out previously by this author (Greene, 2011) that revolved around the extended comparative study of network architecture size where a number of networks were developed and consequently compared on the basis of connectionist model predictive capacity, with numbers of neurons within the hidden layer ranging from 1 to as many as 200. The results were rather promising and naturally suggested that large model sizes are able to provide increasingly better predictive capacity as compared with traditionally employed methods such as logistic regression and systematically selected connectionist networks with simpler network architectures. The process to carry out this type of a comparative study was rather tedious and required not only a lot of time to program the coding for the connectionist modelling, but also was very computationally demanding with modelling process running for weeks non-stop. This approach of course readily identifies an issue with scaling possibility, as with larger networks the training and testing time would be expected to increase exponentially. Another concern is methodological, as the models tested and compared are nevertheless selected by the researcher – together with the scaling issue where connectionist networks with complex model architecture that incorporates multiple hidden layers would pose a serious obstacle that would be extremely difficult to circumvent, and instead would most likely simply remain as an effective limitation of the approach. Moreover, it is important to consider using simulated synthetic data for the methodological testing to determine the optimal approach and architecture as an additional research stage before the actual investigation of the data is carried out to make sure there is no bias in approach selection that is an artefact of the data itself, which is then used to study this very same data during the main stage of the experimental research project.

Here the approach to evaluate the model capacity was developed and carried out quite in the manner as proposed, and simulated data was used to assess the effectiveness and efficiency of the pruning method employed to simplify and optimise the network architecture before developing and retraining connectionist networks using the actual consumer data.

6.4.4 Pruning connectionist models: simulated data

To address the concerns expressed above while assessing the connectionist network performance capacity as part of the comparative analyses that aim to determine the optimal network architecture, before the consumer behaviour dataset is examined, the models are compared and evaluated using the simulated synthetic dataset instead. This should circumvent at least some of the most commonly encountered points of concern as covered in previous paragraphs while discussing and critiquing research design of some of the previous studies.

The overall purpose of this testing and evaluation stage using simulated data is twofold. On the one hand, it is essential to carry out a proper empirical analysis to test and assess the performance of the new pruning capacity that was developed as part of this research project in the statistical package *RSNNS,* which is now available for any researcher to use through the statistical programming language and environment *R,* before we employ these techniques with consumer data here. On the other hand, since the structure and form of simulated data is notably less complex than the consumer behaviour dataset which will be used in the subsequent analyses, it would make sense to carry out some of the analyses on the simulated data initially – this should not make any difference semantically since these preliminary analyses and the results deal for the most part with technical and applied systematic aspects of research design, and therefore the conclusions and learning they are able to offer should be general enough to be applicable to any connectionist model irrespective of connectionist network architecture complexity or the datasets employed.

A number of functions were identified as described in the results section above to test the effectiveness of pruning functionality, and carried out in a manner that would provide sufficient levels of validity and reliability through replications and randomisations. Essentially the test was constructed to allow for a connectionist network to develop an architecture that would be largely excessive for the task at hand, as it could be expected that the connectionist network during the training stages would tend to use all available computational neurons and connections to build a best possible network within the constraints which are set, irrespective of network architecture complexity or size – this of course would essentially be reflected in higher computational demand and network learning time as a result. This also makes it difficult to examine and assess the variable contribution values to assess which of the input variables carry the highest levels of predictive or explanatory capacity in relation to the model output. Whereas pruning algorithm should account for all these factors and systematically force the connectionist network to develop a concise essential network architecture, removing in the process inessential connections weights, which can even result in isolating some of the computational nodes and even bypassing some of the hidden layers altogether. It is essential of course to maintain a sufficiently high level of predictive and explanatory capacity not to sacrifice large proportions of modelling ability in a trade-off for the optimised structure – to address this, the RMSE levels of all models would be consistently recorded.

When the test were carried out, it is apparent that the results were entirely as expected: the connectionist network that otherwise would be rather large and use all available computation and network architecture resources even to maximise the network capacity to the fullest extent, would be substantially trimmed down by the pruning algorithm which would be extremely successful at removing the inessential connection weights to optimise the network architecture – all while the connectionist network predictive and explanatory capacity remained substantially high and uncompromised by the network structure optimisation efforts. Moreover, the pruning algorithm was not only able to effectively remove the inessential connection weights, it also successfully nullified multiple hidden layers – essentially all superficial connectionist network architecture that was not essential for the task at hand was pruned out to leave the bare-bone architecture required at the very minimum to solve the problem.

6.4.4.1 Pruning connectionist models with *Optimal Brain Surgeon* algorithm

The test of course was merely designed to assess the capacity of coding and method of using the pruning algorithm in the statistical programming package *RSNNS* – and was not designed to test the pruning algorithm itself, which would be way beyond the scope of this research project (this would inevitably take research direction towards the field of machine learning – something that could be potentially explored in collaboration as part of the future work). As already mentioned above, the pruning algorithm used through the research project here is the *Optimal Brain Surgeon* (Hassibi & Stork, 1993; Hassibi, Stork, Wolff, & Watanabe, 1994; Hassibi et al., 1993) – the very positive results with pruning and optimisation of the connectionist model structure using the simulated data as discussed above could be largely credited to the sophisticated and elegant design of this pruning algorithm. Hassibi, Stork, and Wolff set out to investigate the use of information available from the second order derivatives of the error function to prune the network architecture by removing unessential connection weights from the trained connectionist model in attempt to optimise and simplify the network to reduce the computational demands, reduce the training and retraining time, and – more importantly – improve generalisation capacity and even further develop the network ability to extract patterns from the data. In the same manner as contemplated here, Hassibi, Stork, and Wolff embarked upon a central problem of pattern recognition and machine learning that revolves largely around the notion of minimising the system complexity, and could often be seen as a problem of regularisation in connectionist modelling: without an appropriate mechanism to minimise a number of connection weights, neural network models could either be prone to overfitting and poor generalisation as a result; or on the contrary could be unable to learn the dataset in an adequate manner if the number of connection weights is insufficient. It is then common to proceed initially with training a sufficiently large connectionist network to a minimum error, and eliminate the inessential weights in a systematic manner to the point where the neural network architecture is optimal – this is where the pruning algorithms specify which connection weights are to be eliminated and the remaining weights are to be adjusted for best performance in the most computationally efficient manner. It was uncovered that *Optimal Brain Damage* and magnitude-based methods have a tendency to eliminate crucial weights –

something that *Optimal Brain Surgeon* never does and is able to maintain a perfect level of performance after pruning is carried out. As a result, *Optimal Brain Surgeon* algorithm is shown to be vastly superior to other magnitude-based methods and *Optimal Brain Damage* (LeCun et al., 1989) – some of which were discussed and critiqued above: for the same training set error, *Optimal Brain Surgeon* algorithm permits pruning of more connection weights than other methods which also often end up removing the wrong weights, thus producing better results with generalisation of test data. Method does not require subsequent retraining – a typically slow cycle after pruning is carried out with other algorithms.

Employing *Optimal Brain Surgeon* pruning algorithm here with simulated data offered very promising results and showed excellent performance in connectionist network optimisation to improve the clarity of the network architecture, which should facilitate the development of exploratory and interpretative accounts with consumer behaviour data. Moreover, the

6.4.5 Pruning connectionist models: consumer data

Having ascertained that the work to implement the coding in the *RSNNS* performs as it should to enable pruning facility, and the *Optimal Brain Surgeon* algorithm is able to deliver the positive result to optimise the connectionist network architecture with simulated data, the second stage is to develop the connectionist models using the consumer dataset.

Considering that the most commonly employed traditional approach in marketing research is a logistic regression, any type of a connectionist network that incorporates hidden layers could be considered a more advanced method: as already discussed earlier, the simplest connectionist network with no hidden layers shows identical level of performance as a logistic regression. Introduction of a hidden layer within a connectionist model opens an entirely different level of performance and capacity. Simulated data was modelled using a connectionist network with multiple hidden layers, but with just a few neurons within each – as was demonstrated, the type of data and a problem only required a single hidden layer with 2 nodes to solve, thus the rest of the network architecture was superficial and therefore was expected to be pruned out by the algorithm. The pruning algorithm performed very well by removing all but the core network architecture required to solve the problem, which was then attempted with actual consumer behaviour dataset rather than the simulated data.

6.4.5.1 Optimal network architecture size

Consumer dataset from Kantar World Panel is of course a lot more complex and includes a number of variables that operationalise transactional, household, and product attributes to describe the purchasing situation and decision-making environment in a comprehensive manner – after the data was normalised and dummy coded, the variables that were selected for the modelling stages that followed ended up being represented by 54 input variables as a result. Keeping in mind that in the traditional marketing literature it is perfectly acceptable and is common practice indeed to study the relations within the consumer behaviour data with logistic regressions – a method that does not have any sufficient capacity to account for the complex relationships that hidden layers are able to capture, and even the available capacity to explore the interactive variables is rarely considered for the reasons that make the process of developing and testing the models exceptionally tedious and poorly scalable – it should be safe to assume that a connectionist network with a single hidden layer with only a few hidden neurons would seem to be able to provide a substantially more advanced model architecture as a result. Thus, 54-2-1 network architecture should be considered to possess a sufficient enough level of complexity to warrant the use of connectionist framework to explore the aspects of informational and utilitarian reinforcement as emergent properties within a hidden layer. Before this claim can be argued however, at this stage it is important to establish that 54-2-1 or architecture of similar level of complexity is indeed sufficient to provide a level of model functionality necessary to satisfy the minimum conditions required for the emergent properties of information and utilitarian reinforcement phenomena to occur.

One of the major critiques of the traditionally employed methods of analysis commonly used to study complex social phenomena such as the act and process of consumer decision-making as argued above is of course the necessity to specify the model architecture and framework by selecting the input variables for the modelling – or just as importantly choosing to *not* select certain variables or not produce interactive variables. Optimally, all possible variables and interactions of all levels and combinations are considered and analysed, and those that do carry the predictive and explanatory capacity are omitted – this is something that traditionally employed methods of analysis such as logistic regression cannot do, whereas something that connectionist models such as neural networks inherently possess as an inseparable part of the computation algorithm, and therefore excel at carrying out each and every time. For that reason, building upon the previous research work carried out by the author and following the research design of others as discussed above, the next experiment was devised as follows: a sufficiently large network architecture is developed to make sure it could train and learn the relations within the data freely, and subsequently pruned to optimise and expose only the core essential network architecture removing all unessential connections. If it is assumed that 54-2-1 network should essentially be sufficiently complex to allow the examinations of the emergent properties of informational and utilitarian reinforcement, the 54-8-8-8-1 network architecture was first examined to see if it would provide sufficiently large and excessive levels of complexity. And it did – in fact, almost too excessive if nothing else: the 54-8-8-8-1 network generates 568 connections between 79 neurons. When pruning is introduced with a few hundred retraining cycles, the model is optimised down to only 66 connections from the original 568 suggesting that perhaps the 54-8-8-8-1 network architecture could be trimmed down quite a bit for all consecutive analyses. In summary however, it is important not to overlook the very successful application of the optimisation algorithm that is able to provide substantial simplification the network architecture, and connectionist methods of training and subsequently pruning the network appear to be a particularly fitting approach to develop the model of consumer behaviour.

6.4.5.2 Pruning different types of connectionist architectures

In the next set of experimental exercises, a few examples of various network architectures are explored. For reasons of simplicity and to optimise the modelling time required – given the lessons learned in the previous set of experimental work that a leaner network architecture would be able to provide a sufficiently complex structure nevertheless, and the fact that each neural network model takes quite a bit of real time to calculate even using a high powered machine with one of the best CPUs available on the mass market at the time, and the fact that to improve the reliability and validity of experimental work these models were replicated hundreds of times – a simpler network architecture (excessively robust nevertheless) was selected that would comprise a more manageable number of 12 hidden neurons distributed among the 3 hidden layers.

Interestingly enough, one of the first connectionist model with 54-4-4-4-1 network architecture was optimised with a pruning algorithm down to a model structure with 54-2-1-1-1 neurons – effectively supporting the initial argument that 54-2-1 network structure is in fact the core architecture that is required to model the relations within the data that represent the consumer purchasing decision-making. Even though this demonstrates that at least in some cases the model naturally removes all but 2 hidden neurons, this was not the most common final optimised network architecture that was observed amongst hundreds of retest connectionist models. In fact, it was observed that with a starting 54-4-4-4-1 network architecture, pruning algorithm was most likely to remove only a few neurons out of the 12 initially available: often a single neuron in one or multiple hidden layers, if any neurons were removed at all. Every time pruning algorithm was able to remove a substantial number of connections irrespective if this would result in pruning the hidden neurons at the same time as well: while a fully connected 54-4-4-4-1 network would contain 67 neurons and 252 connections, pruning algorithm would be able to optimise network architecture down to a much more manageable number of connections, with as few as 67 connections remaining in some pruned connectionist networks.

As it would seem that the shape of the network architecture is able to exert a certain level of influence, it would make sense to test a few different shapes as well to examine what effect this would hold over the performance of the connectionist model. Thus, in addition to the 54-4-4-4-1 network architecture, 2 more types are examined: a funnel-type network architecture with a 54-6-4-2-1 design which could essentially represent a more complex version of the 54-2-1 design; and a reverse version of the funnel that would compress the connection into 2 nodes initially and then allow to grow the network again, with a design of 54-2-4-6-1. First, consider a network architecture with a 54-6-4-2-1 funnel-type design: even though the number of hidden neurons compared with the 54-4-4-4-1 design, number of connections in the initial network before pruning naturally was substantially higher at 358 (as opposed to 252 in 54-4-4-4-1 network) due to the fact that now the many input units were immediately connected to a total of 6 neurons within the first hidden layer rather than 4 neurons as was the case with the models of the previous 54-4-4-4-1 design. This time, starting with a much larger initial network architecture, pruning algorithms was able to trim it down to 78 neurons in the best-case scenario. Even though the absolute number of connections that remained after pruning with the 54-6-4-2-1 network architecture was higher than with the 54-4-4-4-1 network architecture, the pruning efficiency improved substantially. The network architecture that followed the reverse-funnel type design with 54-2-4-6-1 naturally produced lower number of connections initially at 146 total due to having only 2 hidden neurons in the first hidden layer to which the input neurons could connect. Pruning algorithm however was able to eliminate more connections than with the other 2 architecture designs, leaving only 25 connections in the best-case scenario. This means that both network architecture types 54-6-4-2-1 and 54-2-4-6-1 were able to achieve higher pruning efficiency than the original 54-4-4-4-1 network architecture design. It became apparent however that network architectures with a funnel type design almost never removed neurons entirely during the pruning stage; and connectionist networks architectures with a reverse funnel type networks, once compressed to only 2 neurons in the hidden layer, almost never used all the neurons in the second and third hidden layer, usually pruning out around half of them entirely. Therefore it would suffice to propose here that for exploratory and interpretative purposes it would seem that the optimal connectionist network architecture design would be of a funnel type: either a relatively simple 54-2-1 design which provides the best possible level of clarity and should be easiest to use for explanatory and interpretative purposes; or something more complex such as 54-4-2-1 or even 54-6-4-2-1 as examined here to develop a more advanced connectionist model of consumer decision-making process that would be able to extract higher number of patterns and microfeatures from the data, but would of course be more difficult to interpret.

6.4.5.3 Adaptive model learning strategy

Important to note here that the models with 54-6-4-2-1 initial architecture design type systematically did not prune out any neurons – unlike the models with other 2 initial architecture design types, even though the efficiency of removing the number of connections was comparatively high across all initial architecture design types. This may suggest that depending on the available resources within the network architecture, connectionist models are able to adopt different learning strategies, and therefore are able to prioritise identification and extraction of different patterns and microfeatures as a result of this selection. It should be obvious that network architecture design would play an important role in the future research, and perhaps would be an interesting research topic in its own right. This is similar to the results reported elsewhere (for example Cleeremans et al., 1989), and may be a very promising line of inquiry to investigate within the dimension of artificial learning, and what implications this may carry for the field of artificial intelligence in general.

6.4.6 Concise explanatory connectionist model

It should be apparent that it is a particularly challenging task to identify a straightforward way to explain and interpret consumer behaviour. On the one hand, simpler models such as traditionally employed in marketing literature logistic regressions are easier to interpret, yet they are arguably not robust enough to capture the complexity of behaviour – in other words, the explanation may be easy because there isn't much the model is able to offer that needs to be explained really. On the other hand, sophisticated models are robust enough to capture the complex relations within the data and extract the patterns that may offer means to explain behaviour, but at the same time they are not easy to interpret – to some extent, because we tend to employ decomposition as a method to simplify the complex phenomena and make them easier to comprehend, which of course would not be an option here because this very complexity is what makes the models robust in the first place. As a compromise, it would seem to be a good option to use a connectionist model with a single hidden layer that would contain only a few hidden neurons – this way the relations between the input neurons and the hidden neurons are clear and quantified with connection weights, and the hidden neurons could be interpreted as an emergent properties that represent an intricate combination of all the relevant microfeatures extracted from the input variables, and therefore can be treated in a similar manner as the concept of utilitarian and informational reinforcement as proposed here following the connectionist method of analysis and interpretation – in actuality, these concepts are most likely to be too complex to make the clear identification possible, but this is probably as close as it would get to a robust model of consumer behaviour. Once the theory is sufficiently developed to provide the plausible interpretative account using these emergent properties located within the hidden layer, it may be possible to consider connectionist models of higher complexity with multiple hidden layers to explore a more complex mechanism of extracting the microfeatures and patterns from the data.

6.4.7 Predictive connectionist model and pruning

It should be clear that pruning connectionist models which are primarily developed for explanatory purposes offers a range of benefits such as simplified architecture and exposed core relations within the data that can reduce the complexity of interpretation. When it comes to predictive capacity however, the answer is not entirely straightforward as was shown above using consumer dataset: on the one hand, connectionist models with pruning show lower RMSE figures than models without pruning; on the other hand, RMSE figures that can be achieved by connectionist models with pruning are only slightly lower and are very much comparable to the figures achievable by connectionist models that do not employ pruning algorithms. It must be noted however that it is the explanatory capacity of connectionist modelling which is of primary importance in this research project, and predictive capacity is used as a secondary characteristic to assess performance level of modelling as a benchmark. Thus, it is safe to argue that indeed connectionist modelling with pruning algorithms is able to simplify substantially the network architecture optimising it for explanatory and interpretative purposes, while maintaining comparable level of predictive performance benchmarked against connectionist models that do not employ pruning algorithms. Still, if predictive capacity was a primary objective, it should be possible to explore to what extent pruning algorithms can be optimised with aim to improve overall predictive capacity of the model, perhaps holding explanatory capacity as a secondary measure to provide some sort of constraint to make the modelling relevant for a particular behaviour and context; or indeed employ artificial synthesised dataset where relations are known and defined to assess predictive capacity to the fullest extent – this however would have to be to be explored elsewhere. Thus, for a balanced exploratory capacity with a relatively high predictive performance, a connectionist network that employs pruning over a simplified 54-2-1 network could be identified as an optimal model to produce an interpretative account of consumer behaviour in a given context.

6.4.8 Pruning network architecture for interpretation

Even though it is not within the scope of this research project to go as far as develop specific algorithms that could adopt pruning for interpretative purposes specifically, perhaps it would be useful to illustrate with a few examples how useful pruning can really be while attempting to provide an interpretative account of behaviour. For illustration purposes, two types of network architecture are developed and interpreted following on the discussions above: a large connectionist network with 3 hidden layers that does not employ pruning, and the same connectionist network that takes full advantage of pruning algorithms.

Using the same data subset as above for Wales and West that contains 13787 cases, a fully connected 54-4-4-4-1 network is developed. To assure the learning procedure is sufficient to explore the dataset, 100 000 iterations are used here to develop the models. For a 54-4-4-4-1 fully connected connectionist network it takes 4643 sec when pruning is not involved, being able to achieve RMSE of 0.90 as a result. Yet, in a fully connected network with 252 total number of connections (as shown in Figure 31), it is difficult to identify any patterns or draw any conclusions by examining the network architecture – indeed, additional statistical analyses and possibly even adaptive algorithms would be required to examine variable contribution in an efficient manner, which can be used for interpretation of consumer behaviour.

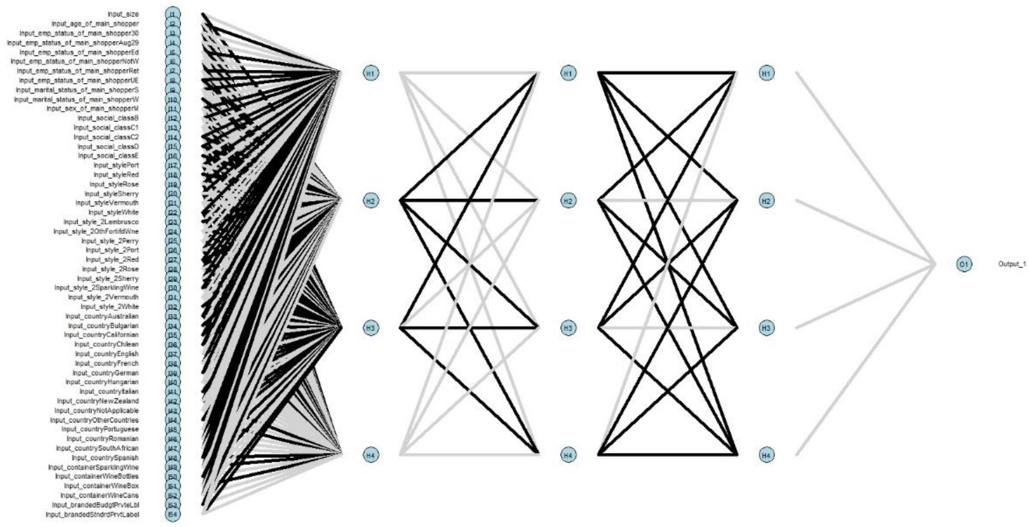

Figure 31. Connectionist network 54-4-4-4-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 100 000 iterations, no pruning.

Whereas when pruning algorithms are employed, it is quite the opposite situation: it takes 5355 sec to achieve a comparable RMSE of 0.95, and produces a substantially optimised connectionist architecture that now only has 51 connections. As shown in Figure 32, pruning algorithms was able to effectively remove 2 entire layers which ultimately were not necessary to model the data and illustrate the pattern within the data used here.

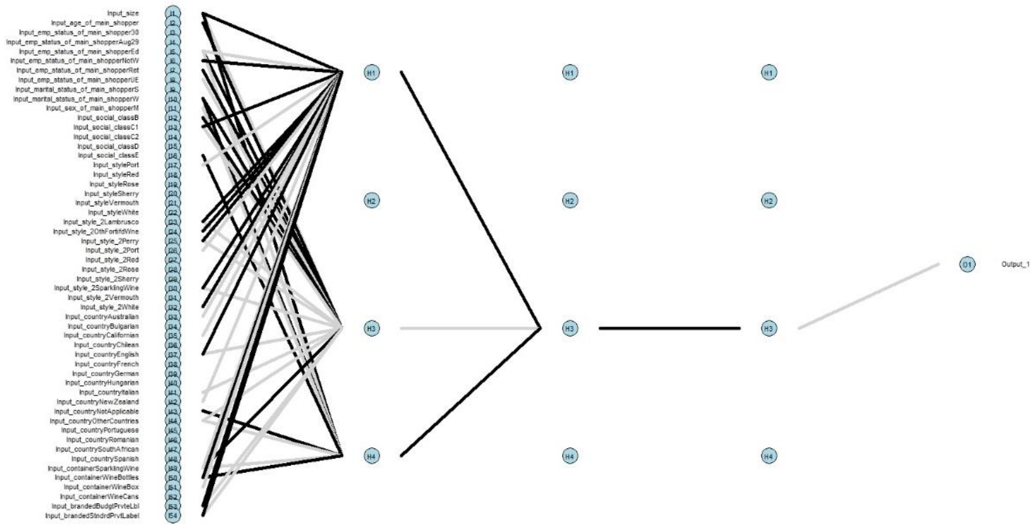

Figure 32. Connectionist network 54-4-4-4-1 architecture using consumer data with 3 hidden layers and 4 neurons each, 100 000 iterations, pruning with 100 retrain cycles.

Upon closer examination, it could be clearly seen that connectionist network that employed pruning was able to establish during the learning process 3 emergent features, which are represented by the 3 computational units within the first hidden layer: 2 excitatory and 1 inhibitory units. Moreover, it is possible to examine which input units contribute towards each of the hidden units which would help to explain what the emergent properties represented by the hidden units may signify, and in what manner in particular. For example as shown in Figure 32, input units for *private label wines* have excitatory connections with hidden unit H1 and inhibitory connection with hidden unit H3, input unit for *other country of origin* (normally cheaper wines) also has an inhibitory connection with H3, and input unit for *age* has an excitatory connection with H3 – this may suggest for example that hidden unit H3 may be interpreted to represent a type of Informational Reinforcement; whereas hidden unit H1 which has excitatory connections with for example inputs units for *not in work* (employment status) *social class C1* and certain types of wine that offer additional benefits such as *sparkling* or *fortified wine* may be very well interpreted to represent a type of Utilitarian Reinforcement. In addition to the illustration of the inhibitory or excitatory connection, connection weights for every connection are also readily available of course for a more precise examination, which can also be used with advanced variable contribution analysis algorithms developed to take advantage of pruned connectionist network output.

It is obvious here that 3 hidden layers are not in fact necessary for this data since pruning effectively nullified 2 hidden layers entirely, thus corroborating one of the previous recommendations that even a simplified neural network with a single hidden layer (for example with 54-2-1 network architecture) could in fact be an extremely powerful method of computational analysis due to its inherent capacity to examine all possible interactions within the data as part of the learning process.

This effectively illustrates the ability of the connectionist network to identify the relevant microfeatures within the data while going through thousands of iterations and exploring any possible interactions between variables, which then can be said to emerge as a representation of higher order faculties that can effectively represent elements of consumer psychology which are used in the decision-making process, or perhaps even be interpreted as a proxy for rule-governed behaviour established as a result of thousands of iterations that can be used as proxy for consumer experiences.

6.5 Theoretical implications

Once convinced that evidence presented above provides a substantiated argument that connectionist models offer extensive predictive capacity in modelling consumer behaviour, it is important to consider what level of explanation and interpretation they are able to provide, and discuss the theoretical implications as well.

It should be apparent at this stage that it is no easy task to produce a truly comprehensive robust explanatory interpretative model of consumer decision-making behaviour. On the one hand, the traditionally employed methods such as logistic regression offer an easier solution open to interpretation – but it is their inherent limitation that makes it impossible to actually capture those truly complex relations within the data that makes the interpretation easier, as the simpler model is unable to capture the complex parts. On the other hand, the connectionist models are inherently designed in a manner that explores all possible combinations and interactions within the data, and through many training iterations learns the data by extracting the complex relations – the process makes it possible to produce robust and comprehensive models capable of capturing accurately the complex relations within the data, but at the same time this very complexity makes it difficult to interpret the results obtained in such a manner. There is just no straightforward way to describe a complex phenomenon such as consumer decision-making process in simple terms without inevitably losing part of the explanation, meaning, or interpretation – simpler terms would ultimately only refer to incomplete and simpler concepts.

Having said that however, it does not mean that there will not be a method to do so in the future. The work carried out here is largely theory developing, where a novel approach is proposed to examine and extend a well-defined and established concept of BPM with a connectionist approach. It is argued and supported with empirical work that connectionist models are well suited to extend the theoretical framework of BPM by providing an empirical evidence to some of its claims and propositions, and proposing structures and processes to continue developing the field further. Indeed, what the author asks is a shift in the level of understanding – from traditional view where input variables which are selected to operationalise a certain phenomenon are expected to link directly into the output variable possibly with a few interactions along the way; to a connectionist fully linked view where input variables are decomposed by the modelling process into microfeatures, which are in turn combined and reassembled in the best possible manner during the training process to form complex patterns that describe relations within the data. These emergent patterns when extracted from the data can then be called the underlying factors that explain the behaviour – this is where concepts like informational and utilitarian reinforcement can be expected to reside. In this sense, informational and utilitarian reinforcement are the theoretical placeholders, as practically it may not be entirely possible to separate the two, as they seemingly tend to form a uniform info-utilitarian reinforcement in most observed cases – the temporarily termed entities to make it possible to develop the explanatory account of consumer behaviour. These patterns very well may describe the utilitarian and informational reinforcement to some sufficient extent, but will also most likely be something a lot more than that, something that is difficult to describe or name because it is a concept that is difficult to comprehend for a human mind: too many variables and interactions are considered simultaneously, and there is no other analogous simpler concept that could help the comprehension. Thus, connectionism is able to provide a theoretical framework and structure, and develop an empirical highly predictive model, but it does not provide any feasible way to decompose the phenomenon into smaller pieces that may be easier to comprehend – simply because this very complexity and the interconnected essence is what we are attempting to explain and interpret.

6.6 Summary

In this chapter, the research findings are discussed and interpreted. First, the concept of informational and utilitarian reinforcement was revisited in connectionist terms, followed by a discussion of results from the exploratory modelling and comparative tests of regression and neural networks. Next, connectionist models predictive capacity and the research around the variable contribution analysis were reviewed. Finally, the advanced connectionist models were developed and discussed, concluding with a conversation about the theoretical implications that identified certain areas for future research direction.

7. Critical assessment of the research project

In this chapter, the precision, thoroughness, and contribution of the method employed here will be concisely discussed; and the connectionist approach will be critically reviewed and compared with its closest rival – the cognitive science.

7.1 Connectionism and cognitive science

Even though it is increasingly common to encounter connectionism while attempting to model mental or behavioural phenomena as the emergent processes of interconnected networks of simple units in many scientific fields such as artificial intelligence, cognitive psychology, cognitive science, neuroscience, and philosophy of mind, nevertheless, connectionism cannot be considered a discipline – but rather a set of consistent approaches that span across these many fields. Connectionism emerged as an approach that integrates the symbolic school of thought, and the behaviourist school of thought: each carry their own scientific framework and philosophy. When applied within a context of a specific field, each school of thought provides a paradigm that defines establishes goals to guide research, and provides its own set of assumptions and techniques. Cognitive science proved to be a great collaborative accomplishment over the years while applying the symbolic approach across many fields of inquiry, whereas connectionism is in a position to question the very foundation of cognitive science and its comprising disciplines as connectionist models provide a plausible alternative explanation to symbolic models.

Towards the end of 1970s, the symbolic systems became a method of choice in cognitive science and its two central disciplines – cognitive psychology and artificial intelligence – as behaviourism seemingly became dated. Soon after however it became apparent that systems solely reliant on symbolic representations and operational rules possessed a number of irreconcilable limitations: the rigid inflexible structures tend to be fragile and offer an inadequate solution to model the process of learning or pattern recognition. This served as an opportunity to reintroduce the network models developed years earlier that would rely on sub-symbolic interpretations and connectionist networks comprised of large number of interconnected computational units, emphasising the distributed representation and statistical modelling. Symbolic systems nonetheless have also been progressively developed to exhibit a reasonable level of flexibility and resilience, learning ability, and subtleness – contrary to connectionist models however where statistical methods are employed to extract patterns from the data, symbolic models maintained the use of ordered strings of symbols.

It should be out of question that connectionist models are now an essential part of cognitive psychology and artificial intelligence, and even more so in engineering and other related fields – this has become more apparent in the recent years in particular as advances in technology reduce the limits of computational ability, and offer the additional capacity that large connectionist models may require. Attempts to reconcile the symbolic and connectionist models over the years inadvertently cultivated an environment susceptible to the idea of hybrid models that would be based on the synthesised framework, incorporating the elements of both connectionist and symbolic systems. Nevertheless, it will still take some time until it is the case that the hybrid systems approach could be considered universally accepted – some additional steps are required such as points argued here for example that behaviourist approach could benefit from incorporating the elements of intentionality to form intentional behaviourism. The level of willingness to modify one's research framework depends on a number of factors that are distinctly dissimilar within the two disciplines – such as the overall purpose of modelling for example.

7.1.1 Artificial intelligence

The primary goal of artificial intelligence is to develop an algorithm that would be able to exhibit a level of performance that could be said to act in an intelligence manner – ultimately an *artificial general intelligence* that would be capable of successfully performing any intellectual task that a human being can perform. In fact, considerable efforts are now increasingly directed at simultaneous development and containment of artificial general intelligence, as it is hypothesised that genuine artificial general intelligence would be capable of recursive self-redesign, resulting in an event of intelligence explosion where intelligence growth would be exponential, quickly exceeding intellectual capacity of any human, and eventually exceeding combined intelligence of all humans. Because it is hypothesised that the capacity of this superintelligence may be incomprehensible for a human-level intelligence, the point beyond which events may become unpredictable to human intelligence is commonly referred to as *technological singularity*. On the one hand, an optimistic outcome is possible where superintelligence would create a utopia for all humans using its superior capacity to solve all world problems; on the other hand however, an opposite outcome would be an event of global human extinction – hence the efforts to contain the superintelligence, and even prevent any chance of its emergence altogether.

Not all research in the field of artificial intelligence is devoted to modelling however, and it is essential to understand that contemporary state artificial intelligence is not only a product of intellectual tradition rooted in the interdisciplinary nature of philosophy, cognitive science, and psychology, but also is a fundamental constituent of it. For example, artificial intelligence contemplates and offers a considerable contribution to central themes such as nature of intelligence and knowledge, theoretical framework of knowledge representation, considering whether certain models can be considered artificial or rather a simulation of human cognitive process – pondering upon these and similar questions is an essential part of the artificial intelligence discipline. Furthermore, it should be possible to view computational algorithms as scientific experiments in a traditional sense, where an algorithm is developed and run, and researchers examine the results to subsequently redesign the algorithm and re-run the experiment – in pursuit to determine whether the algorithm can be considered an adequate representation of intelligent behaviour.

Newell and Simon (1976) argued that any system capable of general intelligence would prove upon analysis to be a physical symbol system of sufficient size, which can be further developed to exhibit general intelligence comparable to the extent of general human intelligence appropriate for any real context and adaptive to the environment within the reasonable limits of processing speed and complexity. In subsequent years, researchers in cognitive science and artificial intelligence explored the research field delineated by this hypothesis, adopting a number of essential methodological commitments: (1) symbols and systems of symbols are used to develop a descriptive account of the phenomena; (2) search mechanisms are designed to explore the inferences that symbol systems could potentially support; (3) it is assumed that a properly designed symbol system would be able to provide a complete causal account of intelligence on its own, effectively removing the need for a cognitive architecture; and finally (4) in attempt to explain intelligence by developing working models of it, the field of artificial intelligence could be considered empirical and constructivist. In the same manner as they are used in natural language where it is understood that symbols refer to or reference something other than themselves, the use of symbols in artificial intelligence is extended to represent the reference to all forms of knowledge, intention, and causality within the environment and context of an intelligent entity – working on the premise that symbols and their semantics can be implanted in formal systems in a constructive manner, introducing the notion of a *representation language*. To model intelligence as a computer algorithm, it is essential to be able to formalise a symbolic system: formal systems allow the assessment of such issues as complexity and comprehensiveness, as well as deliberate the structural organisation of knowledge and complex semantic relationships. The issue of *grounding of meaning* however has been seen as a hindrance by both advocates and opponents of symbolic systems in artificial intelligence and cognitive sciences – it queries how symbols can have meaning, and whether traditional artificial intelligence systems that operate on the principle of linking once set of symbols to some other set of symbols would actually have any ability to interpret these symbols in a meaningful manner in the absence of supporting semantics that are normally available to humans from a social context. As a result, the methodology of traditional artificial intelligence systems focused on exploring the pre-interpreted states and their context, which come pre-encoded by the architects of the artificial intelligence system with contexts and semantic meaning and therefore serve as a function of this particular type of interpretation – as a result, such artificial intelligence systems are able to demonstrate very limited capacity to extrapolate new meaningful associations while exploring their environment. Therefore, the most successful applications that tend to abstract away from the social context to capture the core factors of problem solving with pre-interpreted symbol systems nevertheless remain inflexible, unable to demonstrate generalised interpretation capacity, and lack resilience and ability to recover.

As discussed at length throughout this research project, explicit symbol systems are not the only way to represent intelligence – connectionist frameworks provide useful functionality to understand intelligence in a scientific and empirically reproducible manner. Connectionist networks are models of cognition that do not necessarily rely on pre-determined and specifically referenced symbols to describe it, since the knowledge representation in the network model is distributed across the architecture of the network, and it may be difficult – if not entirely impossible – to isolate specific concepts to particular computational neurons and synaptic connection weights within the model, as any part of the model may be instrumental in the representation of various phenomena. Therefore, connectionist networks serve as a challenge to the argument of Newell and Simon (1976), and in addition to symbolic representations provide a new string of research around the concepts of adaptive modelling and learning for the field of artificial intelligence. Because the structure of the connectionist network is formed by the process of learning as much as by the design, it does not require an explicit symbolic model and rather is a result of interaction within its environment. In such a way, connectionist models are recognised for a number of substantial contributions to understanding of intelligence – more importantly within a context of this research project (but not limited to) a plausible model of underlying mechanisms that describe a learning processes and behaviour, from the viewpoints of both artificial intelligence and cognitive neuroscience. This very inherent nature of connectionism that is so distinctively different from the tradition of symbol models in artificial intelligence is precisely the reason why connectionist networks may be particularly suited to address some of the questions that may be outside of the competence of expressive functionality of symbol models – for example the pattern recognition capacity of connectionist models dealing with noisy data, where distributed representation enables a properly trained network to demonstrate performance similar to humans employing extrapolated elements of similarity rather than logical rules in task such as classification of previously unseen data. It is apparent that neither of the approaches is likely to emerge as dominant, and hybrid solutions that incorporate both symbols and network are necessary to develop truly robust models of intelligence and cognition.

7.2 Distributed representation

In machine learning as a model of information processing, using one computational unit to represent one element is the most straightforward method – commonly referred to as *local* representation – is easy to understand and interpret, as the network structure corresponds to the structure of the knowledge it embodies. There are other implementations however – they are more complex, but at the same time offer notable emergent properties unavailable with local representations. In the following paragraphs, a few notable features of distributed representation as an inherent feature of connectionist models will be discussed – it should make it clear that distributed representation is particularly suited to the task of explaining complex phenomena such as consumer purchasing decision-making process.

7.2.1 Memory

Thinking about consumer decision-making process, the standard metaphor for a memory system is typically a hypothetical warehouse for mental copies of items with some sort of storage and retrieval facilities that find a copy of an item using descriptors provided – this process however is inefficient and contradictory to the process of human memory where acceptable results could be produce even with missing or incorrect descriptions. Another way to view memory is not as a traditional content-addressable search mechanism using available descriptors, but rather as an inferential process: the memory is retrieved by constructing every time a pattern of activity using microfeatures and their connections to represent the most plausible concept consistent with the available cues. Connectionist network models are inherently based on these principles and therefore would be particularly suited to handle elements that are responsible for memory storage and retrieval as part of the consumer decision-making process.

7.2.2 Generalisation

Constructionist concept of memory is related to another feature of consumer decision-making process – the instances when the new items are learned and subsequently stored in memory. To accommodate this in a connectionist model, while at the same time making sure the existing items are not deleted, many connection weights could be adjusted a little – this would have a transferred effect for all related items, while ignoring the unrelated items. This type of model behaviour epitomises the concept of distributed representation, and while doing so invokes the most remarkable of properties – *generalisation*. Rather than using local representations, a distributed representation system would store information by automatically extracting and decomposing the phenomena into its constituent microfeatures, where specific microfeatures could relate to multiple phenomena simultaneously. When new information is acquired, it is automatically propagated throughout the distributed representation system to modify activation patterns for all related patterns, thus making it available throughout the system in a similar manner humans are able to make generalisations.

7.2.3 Learning history and behaviour continuity

In attempting to formulate a plausible explanatory account of consumer behaviour, being able to attach arbitrary descriptions to units using local representation may seem more intuitive and therefore be considered an apparent advantage over distributed representation systems. There is a matter of efficiency that distributed representation systems offer however, which not only allows assigning particular descriptions to the distributed representation clusters in the similar manner as in the local representation systems, but also makes it possible to construct new concepts – a feature that can clearly be useful to facilitate the future development of theoretical frameworks for such concepts as learning history and continuity of behaviour.

7.3 The implications for consumer behaviour

It can be said without a doubt that consumer decision-making process is a complex phenomenon that can normally be attributed to intelligent behaviour – in fact, it could serve as a reasonable test to assess a level of artificial general intelligence: if an artificial system (likely embodied with the use of robotics) could just go out at any unspecified established location and purchase a few required items, all on its own while learning the environment and making other decisions as necessary to reach the final purchase goal, it could be said to act in an intelligent manner comparable to that of a human consumer. Thus, working to develop plausible models of consumer behaviour could not only serve to satisfy the immediate questions that deal with the purchasing decision-making process per se, but also serve as a substantial contribution to modelling cognition and intelligence.

7.3.1 Utilitarian and informational reinforcement, and NNs

In the course of this and previous research projects, it became apparent that a method to empirically define and measure informational and utilitarian reinforcement within the data would be extremely helpful (Greene, 2011). One of the reasons of course is that previous research required a substantial amount of work to consider and define the level of utilitarian and informational reinforcement – for each of the brands within the data. This requires not only an extensive familiarity with the market situation, but also a significant amount of time and resources: if at all possible, a qualitative study is carried out to identify and validate brand perception attributes, which are useful to operationalise utilitarian and informational reinforcement in a subsequent quantitative survey research. Then, traditionally a quantitative investigation would be carried out which would aim to survey a number of individuals and collect brand perception data where each brand is evaluated by respondent with a standardised questionnaire, and with a substantial sample size, both utilitarian and informational reinforcement values can then be attributed to brands to be used in all consecutive research.

To streamline and optimise this process, it was hypothesised that this task could be carried out by a connectionist model that should be able to extract the utilitarian and informational reinforcement for each brand from the data during the pattern recognition (learning) process and subsequently provide a method to quantify them by assigning a numerical score to each brand. Instead, it became apparent that utilitarian and informational reinforcement in fact belong to a qualitatively different level of explanation, and could be modelled in a very viable manner as emergent entities using distributed representation of hidden layers in neural networks (Greene, 2011).

7.3.1.1 Consumer behaviour modelling process

The inherent nature of a neural network as a method of analysis makes it qualitatively different from traditionally employed methods such as logistic regression. For the purposes of this discussion, let us pay no attention to the methodological differences and instead consider the semantics. Neural network model with no hidden layers essentially is no different from a logistic regression as it too only incorporates the input and the output layers. Once the hidden layers are introduced between the input and output layers, the neural network is capable to develop unique features that are of particular interest in the discussion of behaviour analysis.

In the case of regression, the analysis only considers pre-specified by the researcher input and output variables, whereas a neural network learns the structure as part of the process. Complex neural networks that incorporate hidden layers determine connection weights between the variables and the hidden layers, depending on the network architecture. The key difference with the traditional method is the meaning of hidden layers and neurons that emerge from the learning process – network characteristics that are not defined by a researcher but rather are the product of a learning process. In the context of consumer behaviour and using the established framework of BPM for explanatory purposes, it is the central hypothesis of this research project that the hidden layers may be interpreted as emergent representation of utilitarian and informational reinforcement, along with other factors that may influence behaviour otherwise inconceivable to the researcher as identified by the neural network in the process of learning and extracting the patterns from the data. What is argued here then is that within the modelling of consumer decision-making process, the utilitarian and informational reinforcement should be positioned on a different semantic hierarchical level to the traditionally employed inputs such as product parameters, consumer demographics, decision environmental parameters, etc.

Thus, it makes it possible for the product characteristics and demographics to be connected not solely to the output variables directly, but rather to intermediary abstract entities following the distributed representation principle, which are shaped by the learning process of the neural network. These abstract characteristics are not easily interpretable however, and it may in fact prove quite difficult – if not entirely impossible – to assert reliably whether the constructs represented by the hidden neurons and connection weights truly represent utilitarian and informational reinforcement. Within the theoretical framework of BPM, utilitarian and informational reinforcement is defined in terms of money allocation as a function of brand reinforcing attributes (Oliveira-Castro, Foxall, & Wells, 2010). This issue however is not absent elsewhere either, as it is not possible to say to what degree the utilitarian reinforcement could truly be represented for example by the additional desirable product attributes such as sausage in baked beans versus plain baked beans (Oliveira-Castro et al., 2010). This assumption may not holds for all consumers, or even may very well be the opposite case for some – consumers may receive higher utilitarian reinforcement from plain baked beans as for example in their particular situation plain baked beans may be used in a wider variety of dishes for example. Thus, the real issue is the matter of operationalization, and is not unique to neural networks.

Even though it should be reasonably possible to conclude that testing this using a pruned and optimised connected network where input units are connected to hidden layer that would contain two hidden neurons to represent utilitarian and informational reinforcement as proposed here in the previous chapters as the preferred option to reduce the level of difficulty while interpreting the final neural architecture, a number of other network architecture types were explored and presented here in attempt to identify the optimal type most suitable for explanation of consumer behaviour, may that be for predictive or interpretative purposes. The abstract characteristics that emerge during the learning process within the hidden layers however are not easily comprehensible or interpretable for a number of reasons. For once, the network capacity allows exploring the incredibly complex interrelations within the data, and able to identify subtle unexplainable patterns. These patterns however could be too multidimensional to comprehend for a human mind – something a researcher would not be able to think of on their own as would be required in the case of regression analysis where all variables and the structure must be predefined from the onset.

7.4 Cognitive process simulation

Thus far, relatively simple connectionist networks have been discussed here that model isolated cognitive processes with simple input and output. But what about modelling higher cognitive processes that would necessitate complexity in both the processing and the input and output of the network? The two networks discussed next are developed as possible models of higher cognitive processes.

7.4.1 Past-tense acquisition model

Ability to construct an infinite number of grammatically correct sentences is based on grammatical rules used in a natural language. In psycholinguistics, it is assumed that either an innate knowledge regarding the rules of grammar is available or humans follow a process of hypothesis formation and subsequent testing based on the experiences and linguistic information available. The product is assumed to be a mental structure in the human mind that contains representations of linguistic rules. Rumelhart and McClelland (1985b) propose a mental structure capable of processing natural language that does not require explicit representations of the rules. English past tense acquisition is a well-studied phenomenon, where a U-shaped learning course is characterized by the three stages. First, past tense for a small group of verbs (mostly irregular) is acquired. This is followed by the acquisition of past tense for a larger group of verbs (mostly regular) where the rule seems to develop (add *-ed* to verb *stem*) – as a result it is also incorrectly applied to some of the irregular forms. In the final stage, regular and irregular forms are used correctly suggesting that exceptions to the rule are learned. Irregular verbs could be further grouped into subcategories, which could explain some of the errors produced by human learners. It is important to set the limits of what is being modelled to simplify the model considerably yet enabling it to achieve the substantial level of simulation. The model of Rumelhart and McClelland (1985b) consisted of three connected networks, where the first network translated the phonological input into the appropriate format for the second network, the pattern associator. Pattern associator output was again translated by the third network into the final phonological output. Networks are quite different from the traditional methods, and therefore pose certain difficulties for traditionally employed strategies. For example, ordered mapping of phonemes would not work as easily in network architecture – instead, context sensitive phonemes were employed. An issue of network efficiency was solved by replacing the representation of all phonemes with the features of the phonemes, which substantially decreased the number of units necessary to encode the problem. This offered distinctive enough representations yet allowing the degree of generalization for a network to generate past tense forms for previously unseen verbs. Thus, the network functionality is not tied to the verb stems in particular but rather to the distributed phonological features, at the same time identifying similarity between the verbs and determining which of the verbs require the application of the regular past tense rule. Even though this model is able to achieve high levels of performance, one critique argues that much work is done in the featural decomposition of phonemes, which is based on the adaptation of traditional linguistic featural analysis (Pinker & Prince, 1988). Even though this may be true to some extent, the contribution of the connectionist network lies in the pattern recognition process not reliant on rules, which is necessary for learning to occur in the network.

For the simulation, encoded verb stem is supplied to the two-layer feedforward network that employs a stochastic version of logistic activation function. After obtaining the pattern of activation, error correction procedure facilitates learning through the adjustment of connection weights by comparing the obtained pattern with the desired output pattern. This network however is not particularly suitable for the purposes of examining the behaviour in detail due to its size, as it was primarily tasked with simulation of the stage-like learning process of past tense acquisition observable in children (U-shaped learning function). Smaller purpose built networks with fewer weights can be scrutinized to examine the emergence of behaviour and the underlying factors that influence it.

7.4.1.1 Overregularization in a simpler model

To examine the processing that occurs inside the network, a simplified model is considered by McClelland and Rumelhart (1988) that comprises of eight input units and simple enough rule used to transforms the input pattern for the output of 18 cases in total. Predictably, due to a systematic nature of the input-output relation, the network was able to achieve absolute performance without actually relying on the rule. Then, one of the cases was transformed to be in conflict with the rule employed to transform the input patterns. To simulate the children learning process, the network was presented with only two cases: regular (that follows the devised rule) and irregular. The network achieved good learning level, but was not able to extrapolate the rule from just two cases properly. When the remaining 16 cases were introduced, the network was able to extrapolate the transformation rule. At that stage, overregularization could be observed with the one irregular case (signified with the increased errors), which subsequently followed by learning to incorporate the irregular case. Thus, the network learning process closely resembles the stage-like learning process of children described earlier.

7.4.1.2 Past-tense acquisition simulation

The learning simulation of Rumelhart and McClelland (1985b) consisted of stage-like process where ten most frequently encountered verbs were supplied to the network first, followed by a set of verbs of average frequency of usage, which was finally followed by a set of verbs with low frequency of usage. The model was able to achieve high level of performance (between 80 and 85 percent) on the first set of verbs after 10 epochs, when the second set of verbs was introduced. This resulted in a temporary drop of performance (around 10 percent) on irregular verbs – a characteristic feature of Stage 2 in children learning process, a result of interfering with learning the regular pattern. By epoch 20, the performance started to improve, and by epoch 160 the performance of the model was around 95 percent features correct – the model was able to learn the irregular forms as exceptional to the rule cases. Mistakes of the model in the general direction of overregularization, as expected (for example, *–ed* added to the stem of the irregular word to forms such as *comed* or *camed*). Thus, the network is able to simulate the stage-like learning and the effect of overregularization without the use of rules. Testing the model on the previously unseen set of verbs with low frequency of usage showed high level of performance (92 percent correct feature activation for regular and 84 percent for irregular verbs). The ability of the network to generate past-tense form to novel verbs was less than optimal, but comparable to human performance in similar task, suggesting comparable limitation of the model to the native human speaker.

The in-depth analysis revealed that many of the distinct subclasses of regular and irregular verbs (for example, nine subtypes of irregular verbs described by Bybee & Slobin, 1982) could be identified in the simulation results. The simulation showed similar ranking of the performance within each of the subclasses, even though variances were less dramatic. This could be due to the fact that the explanations responsible to performance variations were not present in the simulation, and therefore may provide a superior account of human performance. In addition, it became apparent that the error type propensity (*comed* vs *camed*) differed across verb subclasses.

Based on these findings, Rumelhart and McClelland (1985b) argue that it is possible to simulate the essential characteristics of human learning behaviour with relatively simple network architecture and without the use of explicit rules.

7.4.1.3 The role of input

In past-tense learning, the role of input in both human and network simulation models is not entirely understood. What requires further examination is the comparison of the input conditions of children and simulation models, and the range of conditions under which U-shaped learning can be present in networks.

7.4.1.3.1 The role of input in children

Some researchers argue that supplying discontinuous input for the network is not a proper mechanism to attain a U-shaped learning curve, as there is no sudden change in vocabulary size or verb type and therefore it should be a factor to have an effect on overregularization (Ullman, Pinker, Hollander, Prince, & Rosen, 1989). If not change in input, then what activates Stage 2 learning? Research indicates that during Stage 1 irregular type verbs outnumber regular type (Bloom, 1970; Nelson, 1973). In Stage 1, children produce the stems for most verbs, but also tend to use incorrectly some of the irregular past-tense forms instead of the stem (Kuczaj, 1977). In Stage 2, proficiency of using past tense in appropriate context improves for most verbs, with the exception of about one-third of irregular verbs where overregularization occurs. If the same type of mechanism was responsible for learning in Stages 1 and 2, it would not be able to control the word production initially and only use the verbs as input, gradually developing the network mapping as more verbs (largely regular, as regular verbs surpass irregular by the end of Stage 2 as suggested by Ullman et al., 1989) become available. As the mechanism progressively matures following the developmental process that improves the coordination of previously separate competencies (Bates, Bretherton, & Snyder, 1988), it gradually develops the capacity to control the production of verbs. Therefore, it supports Rumelhart and McClelland's claim that these changes facilitate the transition from Stage 1 to Stage 2. The simplification of the process for the simulation model could be further justifiable in light of lack of understanding how the process actually occurs in children. The two suggested developments could either examine in greater detail the child acquisition data, or the network behaviour.

7.4.1.3.2 The role of input in networks

When the network is trained on a small subset, it is capable of achieving high performance by learning the whole dataset without extracting the patterns. In the case of learning the past-tense verbs, it may seem to be just that, as the initial set of verbs available to a child normally contains a rather small number. Once more verbs paired with past-tense forms become available, the network begins to develop the evidence of a systematic structure in weights. As the systematic structure becomes more pronounced in network weights when more verbs are made available for inputs, the inclination towards overregularization becomes evident as well - even if only half of the input verb pairs exhibit the pattern explicitly.

Plunkett and Marchman (1989) performed a number of past-tense formation simulations where the network was presented with a complete set of verbs at all training epochs, effectively eliminating the input change altogether. Regardless, U-shaped learning was obtained for some individual verbs. They employed a much simpler distributed coding for the verbs as well compared to which was used by Rumelhart and McClelland (1985b) and therefore did not promote generalization as much.

7.4.1.4 Past-tense formation model summary

Rumelhart and McClelland (1985b) argue that relation between the verb stem and past-tense form is described by a set of general rules, but it is the mechanism of distributed processing across the network weights that governs the relation.

Moreover, both standard and exceptional cases are encoded within the same single network architecture, and network learning process shows similarities to the learning process of children.

7.4.2 Kinship knowledge model

In psycholinguistics, the ability to handle kinship relations is a common test to assess the model. (Hinton, 1986) developed a multi-layer connectionist network for such a task, where the information on 24 individuals is analysed by a five-layer network with three hidden layers. Individuals are divided between two families with isomorphic family trees that include 12 relationship types. Inputs contain Person 1 and a relationship type, and the model provides the encoding for Person 2 on the output. The first hidden layer contains 12 units: six units receive input from 24 Person 1 units, and six from 112 relationship type units. The second hidden layer contains 12 units that are fully connected to the first hidden layer, and provide the input for the third hidden layer that contains six hidden units and provides the input for the output layer that specifies the output for Person 2. The network is forced to extract the relevant features for the distributed representation as the input information follows through the layers that contain fewer units (36-12-12-6-24 network structure), and then use the extracted features to identify the output. Out of the 104 possible Person 1 - relationship - Person 2 cases, a back-propagation network was trained on 100 cases and tested on the remaining four cases. The network was able to provide a correct output for all four cases in the first run and for three out of four cases in the second run, suggesting a high performance capacity without the reliance on propositional representations or inferential rules.

When the weights matrix is examined, it becomes apparent in what way was the network able to accomplish the cognitive simulation task. For example, the connection weights between the input units representing Person 1 and the first hidden layer suggest the extraction of features relating to the family symmetry, such as which of the two families or generations (younger or older) the person belongs to. Thus, the network was able to identify the kinship structure only from the cases of specific relationships presented to it through the restructuring of the information and feature extraction procedures. Based on the internal featural distributed representation developed as a result of the training process, the network was able to learn the nature of relationships and use this knowledge to make inferences. It is imperative however to be aware of the hidden unit interpretation, as in most cases, even if it may seem quite natural to assume certain labels from the examination of network behaviour, the hidden units represent a very complex interrelated combined subset of subtle features extracted by the network that may not be straightforwardly explained.

Many questions remain however that revolve around the interpretability of hidden units, and whether the interpretation is even necessary or beneficial; the training and testing procedures; and questions that deal with the type of higher cognitive tasks that can be simulated with networks.

7.5 Phenomenological critique of computational models of intelligence

One type of critique originates in phenomenology, and disputes the established computational approach of modelling intelligence. On the basis that the process of skill acquisition has been misunderstood, Dreyfus (1992) argues against general approach to modelling intelligence which was undiscerningly adopted by early Artificial Intelligence researchers – the same approach that forms a fundamental part of the research programme described here. Building upon the work of Heidegger (for extended discussion please see for example Heidegger, 1988), Dreyfus supports his opposing view with an assertion that humans are in fact experts at carrying out a multitude of tasks within varying situational context as part of everyday life. This type of expertise is arguably overlooked in traditional approach to computational intelligence modelling, and instead is taken to be an assumed foundation upon which all subsequent learning and rule formulation occurs. Dreyfus argues quite the opposite by stating that humans begin with a pre-formulated set of explicit rules, which are applied and specialised to a multitude of particular contextual situations. The key element of critique postulates that computational algorithms do not generate meaning or sense, but rather appear as meaningful when taken within the context of human everyday expertise. Thus, Dreyfus asserts that the understanding of intelligence with the use of computational modelling would inevitably entail revisiting the fundamental principles of meaning that originally identified the need for these computational and technological artefacts, and how it reflects upon the human identity.

Certainly, these arguments may hold true while considering human intelligence, but would they be appropriate in the same manner to explanation of artificial intelligence? In fact, some of the aspects of intelligence that are taken to be virtues with human intelligence, such as human everyday expertise as discussed above, could be seen as a limitation and a constraint when speaking in terms of just any general intelligence that is not tasked primarily with closely replicating the way human intelligence occurs, and instead developed to achieve certain level of performance by optimal means, which does not necessarily need to be similar to that of a human intelligence. Positively, this research direction is further supported by recent technological advancements that constitute a substantial progress already, and computational models are eventually expected to reach performance levels comparable to the most advanced currently known information processing entity, the biological human brain – and indeed surpass not only performance level of individual humans, but ultimately surpass the combined collective intelligence of all humans. Certainly, this would entail inherently different way for intelligence to emerge.

7.6 Summary

In this chapter, the concise assessment was offered to critically assess the connectionist approach against the established tradition of cognitive science.

Some of the inherent features of both disciplines and their modelling approached discussed to evaluate both advantages and disadvantages in relation to the task of explaining the process of consumer decision-making.

The following chapter will consider a few potential direction for future research.

8. Future work

It is clear that even though the experimental work carried out as part of this research project is able to offers a substantial amount of data for all subsequent analyses carried out and discussed here, there is a multitude of unanswered questions that could extent this line on enquiry further, and in a number of potential directions. A few of these potential areas of inquiry will be discussed in this chapter.

As a continuation of the explanatory modelling approach predominantly discussed here, it could also be advantageous to continue developing explanatory capacity by taking it in a number of different directions, some of which are discussed next.

Once an acceptable level of explanatory capacity is achieved, it then becomes possible to explore the prescriptive direction and normative consumer behaviour modelling in attempt to optimise the connectionist modelling and develop a certain level of prescriptive capacity to achieve specific objectives that may be desirable for a number of reasons. Some of these are discussed in the following paragraphs.

8.1 Individual respondent level of behaviour analysis

One obvious option to advance the research undertaken here would be to consider individual respondent level rather than a multi-respondent prototype defining method of analysis carried out here. Obviously this is outside the scope of this research project, but would seem a natural evolution of this line of inquiry, which is further corroborated by some of the other potential directions to advance the work discussed here as proposed and discussed in the next paragraphs.

The problem of demonstrating individual behaviour in a scientific manner is reasonably well understood and comprehensively described (for an extended discussion please see Skinner, 1953), and over the years has been applied to a wide range of contextual and behavioural settings which resulted in producing general descriptive accounts of mechanisms that govern and foster many observable individual forms of behaviour. Thus, the research process that carries out an applied behaviour analysis on an individual respondent or consumer level is a self-monitoring and self-evaluating method of scientific inquiry to study behaviour in experimental applied manner. The pragmatic nature of behaviourist theory is evident in the application of behaviour analysis, where the verbal description of non-verbal behaviour by the respondents themselves would not be acceptable, and the focus of the research programme is in fact revolves around what subjects can be brought to *do* rather than brought to *say* – that is unless verbal behaviour of interest of course.

It would be customary to expect for a consumer behaviour to be composed of a number of a sequences of physical events, and precise measurement of these events is required for a scientific examination – this bring upon the problem of reliable quantification of the behavioural response which cannot be easily circumvented in applied consumer behaviour research, whereas in non-applied consumer and other research there may often be an opportunity to select a behavioural response item which is easier to quantify and measure in a reliable manner from the onset. Behaviour analysis normally requires a demonstration of sequential events that are said to be responsible for the manifestation of the behaviour, and researcher is required to demonstrate an evident degree of control over the said behaviour by either being able to increase or decrease the frequency or duration of behaviour – something which is reasonably achievable within the experimental laboratory setting by either replication or satisfactory probability levels derived from the statistical analyses and modelling of grouped respondent data, yet may pose a difficulty in an applied context of consumer decision-making situation in a market environment. The two types of research design that are commonly used to demonstrate with a certain degree of reliability the behaviour control can be referred to as the *reversal* and the *multiple baseline* methods.

The first method assumes a continuous tracking of behaviour for an extended period of time to ensure the clear measurement stability is achieved before the experimental variable is applied. The behaviour continues to be monitored and measured to determine if the experimental variable is able to exert any significant observable change of behaviour – if it is indeed the case, the experimental variable is discontinued or otherwise altered to examine whether the observed behavioural change is dependent on the experimental variable, as it should result in the observed behavioural change to diminish and reverse (hence the naming convention) to the levels as initially determined in the first stage of the experimental research design. The experimental variable is then reintroduced, yet again to observe if this would recover the behavioural change to the level as previously observed in the second stage of experimental research design. The reversal procedure with a subsequent measurement could be carried out a number of times if the experimental research setting permits the multiple reversal stages to improve the validity and reliability of the obtained results – something that is unlikely to be the case in the applied context of consumer decision-making situation in an actual market environment. In fact, it may be difficult to carry out even a single reversal cycle, particularly when the behavioural change caries a positive measurable commercial impact and one may be reluctant to abandon the favourable results and revert to the original level for the sake of experimental design. In contrast with experimental laboratory setting, the dynamic market environment may also be difficult to control and contaminate the applied research design as naturally occurring changes may be brought upon by other external factors that lie outside the control of the researcher. It may also be unethical to reverse some of the valuable behaviours that are able to demonstrate as a result particularly positive and beneficial effects – for example research programme that aims to improve the consumer purchasing consumption decision-making habits in attempt to decrease the incidence of certain diseases or other serious health risks. Moreover, it should be expected while producing a valuable behaviour in a social setting to generate a degree of extra-experimental reinforcement from the social setting itself; and as a result, the valuable behaviour may cease to be dependent upon the experimental design and behaviour control that was set up to produce this very behavioural response in the first place.

The second method of *multiple baselines* may be particularly useful as an alternative to *reversal* method as it could allow to overcome the difficulties within the applied context of consumer decision-making in an actual market environment and circumvent some of the potential issues described above. To do so, a number of behavioural responses are identified and measured over time to determine the baselines for all consecutive research work. Once these multiple baselines are reliably established, the experimental variable is introduced with one of the identified behaviours – the resulting behavioural change is recorded while other baselines are monitored in parallel to identify any other concurrent change that is not associated with the experimental variable. In the case of success with the first application of experimental variable where the significant observable behavioural change can be identified, rather than discontinuing or altering the first experimental variable to reverse the newly created behavioural change, instead the experimental variable is introduced to one of the other yet unaffected baselines. If the significant behavioural change can be observed again with the second baseline, this would increasingly demonstrate the effectiveness of using the experimental variable to control behavioural response as opposed to change occurring as a natural or random variance – at this point to improve the validity and reliability of the results thus obtained, the experimental design can be systematically extended for the remaining baselines by introducing the experimental variable to yet another baseline at a time while monitoring and tracking baselines for all behavioural response items in parallel. Even though arguably the *multiple baselines* research design is better suited than *reversal* for studying consumer decision-making within the applied context of market environment and allows circumventing some of the potential limitation due to practical and ethical issues, the element of qualitative judgement is still necessary nevertheless to assess the suitability of inferential statistical analysis: for example to determine and specify the number of baselines (the same in the case of reversals) required to provide a convincing and satisfactory account of demonstrating a reliable behavioural control.

These two research designs are of course the core foundations upon which more complex composite and combined designs could be constructed – indeed it may be required to decompose each successful demonstration into its comprising elements to be studied and examined separately, perhaps employing certain variable contribution analyses to determine the extent of contribution of each component to the overall behavioural control; or perhaps introducing additional elements to assess the generality of behavioural change as able to remain durable over time; or perhaps examine the sequence of variables required to improve the generality of behavioural change in the best possible manner. While using secondary data, it may be increasingly difficult or perhaps even impossible to identify the suitable cases where the behavioural change can be observed, but perhaps some of these difficulties could be bypassed by a robust research design and advanced modelling techniques that are inherently based on the principle of individual stand-alone subsystems which are able to exert behavioural change as a matter of collective contribution, such as *swarm intelligence* methods discussed later.

8.2 Multi-category behaviour analysis

Certain practical and ethical limitations pertaining to research design described in the preceding paragraphs above may also be true in attempt to study and model consumer behaviour employing multiple product categories simultaneously and in parallel. The scope of this research project is limited to a single product category – wine, and proposition to examine multiple product categories simultaneously would substantially increase the analytical complexity required, as it would inevitably introduce the need to account for the dimension of interactive cross-category consumer choice. There are a number of other reasons to follow the boundaries set here as well that prescribe the inclusion of a single product category – not the least of them is a high computational requirement that is necessary to process the large amounts of data, which is the case even with a single product category. If the computational power was not an issue where parallel and distributed computing that links multiple processing units or even the use of supercomputers can be employed, it could be beneficial to explore the capacity of connectionist networks to develop cross-category models. On the one hand, connectionist model generalisation capacity could be assessed where connectionist model trained on one category could be tested to predict behaviour with an entirely different product category to carry out a true out-of-sample validation. On the other hand, the connectionist models could be trained on multiple categories from the onset, which could enable the extraction of mode general non-specific to product category patterns that represent consumer behaviour more accurately, in turn making the connectionist models perform better with single and multiple categories as well.

The notion of considering multiple product categories to develop better understanding of behaviour has been reviewed over the years by many authors in the different yet closely related thread of research that revolves around the concept of *market basket choice* (for an extended discussion please refer to Chib, Seetharaman, & Strijnev, 2002; East, Hammond, & Wright, 2007; Manchanda, Ansari, & Gupta, 1999; Russell & Petersen, 2000). Without a doubt, as the technological advancements facilitate the inevitable improvements in data collection by refining the methods and growing the number of participants in the panel data to improve the extent of representative sample and improve the product and behaviour coverage on both longitudinal and individual levels, researchers are increasingly involved in developing statistical modelling techniques of basket-level multi-category consumer decision-making behaviour. Global data providers are now able to offer truly massive integrated databases (such as the complete dataset of Kantar World Panel which only a single category was employed for the research programme described here) that include longitudinal data and information on a household level for a number of observed behavioural transactional variables such as store choice, category incidence, brand choice, and quantity purchased – all this is of course supplemented by an extensive household demographics, product attributes, and purchasing decision environment (store and venue attributes). Indeed, the field of consumer purchasing behaviour and consumer choice modelling spans a number of decades, and over the years a multitude of models and methodologies have been proposed, even some that attempt to develop comprehensive models that cover a number of behavioural variables concurrently (for example an attempt to model simultaneously incidence, brand choice, and purchase quantity by Chintagunta, 1993); the bulk of choice modelling research however has been limited to a research design that considers a single product category at a time – very much like the research programme described here which is deliberately constrained to a single product category, wine. In endeavour to extend the choice modelling research by addressing the limitation of a partial single-category consumer behaviour line of inquiry and providing a plausible account of interdependencies between the multiple product categories, the notion of *market basket choice* not only attempts to model the multiple purchase outcome variables simultaneously, but also across multiple product categories – that is, develop an understanding of choice behaviour across multiple, and ultimately all, product groups that comprise the total shopping basket. The interest to extend the multiple category choice research is not exclusively motivated by reasons of academic research, but also commercially driven, and maybe even to the higher extent: retailers strive to maximise total basket spend, organisation that hold significant presence within multiple product categories increasingly optimise multi brand portfolios and minimise cross-cannibalisation of profit and return, and customer acquisition and cross- or up-selling initiatives require cross-category models to improve targeting and segmentation analyses. Another limitation that contributed to the historical restriction of choice modelling to a single product category that can now possibly be overcome at least to some extent as a result of technological advances is of course the very high computational demand to carry out the multi-product modelling – naturally it is not simply a matter of additive computation burden where multiple categories would result in a linear increase in computational requirements dependent on the number of categories modelled simultaneously, but the computational requirements would rather increase exponentially, as the very point of the efforts to carry out a multi-category choice modelling is to consider all possible inter-category relations within the data. Nevertheless, recent advances in multiple product category modelling have been able to propose a number of approaches as discussed in the following paragraphs.

Multi-category purchasing incidence research surveys the dichotomous consumer decision-making situation by considering the notions of product substitutability and complementarity within a constraint of a limited disposable monetary resource. The research indicates that simultaneously modelling the incidence of higher number of product categories with multivariate probit and logit models may mitigate otherwise overestimated effect of purchasing situation and environment (Chib et al., 2002; Manchanda et al., 1999). Moreover, it is suggested there may be a consistent positive correlation among all product categories, which may be indicative of an inherent inclination for simultaneous incidence across all product categories. Building upon the dichotomous choice models to explore the effect of time-sensitive price elasticity of multi-category purchasing incidence, multivariate additive risk models can be employed (Ma, Seetharaman, & Narasimhan, 2005) to model the purchasing frequency variable. Another research direction emphasises the prominence of underlying product attributes to describe consumer multi-category purchasing behaviour (Chung & Rao, 2003), while studying the bundled multi-category products. It should be apparent that these and other comparable approaches are able to contribute incrementally to the understanding if the multi-category purchasing incidence, and these learnings could potentially be amalgamated to develop a cohesive integrated framework of multi-category purchasing incidence.

Cross-category brand choice models investigate the relevance of marketing mix sensitivity and cross-category brand choice correlations in the context of multi-category purchasing. Employing decomposition methods to differentiate between household-specific and category-specific components of the marketing mix to examine whether the inclination of the household to stability of variety in its brand choice would carry over multiple product categories, Ainslie and Rossi (1998) discover high levels of cross-category correlations suggesting it to be the case indeed, and propose that those household which exhibit strong brand loyalty in their choice behaviour over extended time periods in one category are likely to exhibit similar behaviour across categories and not pursue brand variety in another category. This notion could become rather useful in a situation where the information about consumer behaviour is only available for a particularly product category and is limited or inadequate for other categories – because of the high degree of correlations of household price coefficients across product categories, it is possible to extrapolate and predict the consumer behaviour and purchasing information from a well-understood product category to another related less-understood product category (Iyengar, Ansari, & Gupta, 2003; Singh, Hansen, & Gupta, 2005). Modelling cross-category correlations of household brand volume purchasing shows high consistency in choice of store and some but not all national brands (Russell & Kamakura, 1998), and branded products show high correlation across multiple categories (Erdem, 1998).

In attempt to combine some of these models and develop a more comprehensive approach to cross-category purchasing behaviour, computational requirements would be substantial and require as a result a certain set of theory-driven restriction to be imposed – flexible statistical model with unlimited parameters, a better computational framework is required. Perhaps one way to structure the approach in such a way that it would be able to offer a reasonable account of multiple and cross-category consumer choice is to decompose the system of individual consumer behaviour into multiple sub-systems that would model consumption for a specific product category, while at the same time interacting with other sub-systems that represent the consumer choice for a different category of the same individual consumer, and also interact on a higher level with other composite systems that represent other consumers, and the elements of the consumer decision-making situation in a market environment. The *swarm* *intelligence* methods could perhaps suggest a plausible solution to these concerns, as discussed in the following sections.

8.3 Swarm intelligence methods

The term *swarm intelligence* first introduced by Beni (1993) is commonly referred to a concept that can be described as a collective behaviour of natural or artificial stand-alone self-organized sub-systems or agents. The inspiration for the *artificial smarm intelligence* came from biological systems encountered in nature such as ant colonies and bee hives which typically consist of a population of simple agents interacting with each other and their environment. The individual agents are said to follow simple procedures, and while the centralised control structure that would specify how stand-alone agents should behave collectively is effectually absent, the global collective behaviour that can be said to be intelligent can eventually be observed to emerge as a result of sequences of random local interactions of the stand-alone agents between themselves and the environment. Artificial swarm intelligence systems and algorithms have been adopted to utilise these principles and used in predictive analytics in the context of forecasting tasks.

The universal prosperity of biological swarm systems such as ants, bees, or termites could be attributed to a number of wide-ranging characteristics: (1) the flexibility required to adapt to a changing environment, (2) the collective robustness of the system that diminishes reliance on individual contributing agents and offers the effect of graceful degradation in the worst case scenario, and (3) the principle of self-organisation that eliminates the requirement of a global control and local supervision. These characteristics are general enough to be applicable beyond the world of biological systems, and can be desirable descriptors of systems and organisations not only from the academic research point of view, but could also bring substantial rewards if applied in the industry as well. Without a doubt, there should be no issue with embracing the first two principles almost universally, whereas the third principle may seem counterintuitive to the traditionally established ways of working and hierarchically structured systems in place. Nevertheless, the success of biological systems and the effective pioneering applications of the swarm intelligence principles in the industry should promote the consideration and acceptance, while encouraging growing interest to develop these systems further.

The usefulness of swarm intelligence should not be limited to the obvious application of biological foraging models to inspire specific types of optimisation algorithms that are able to provide superior solutions with large complex tasks that deal with high levels of ambiguity such as a probabilistic optimization algorithm *ant colony optimization* (Dorigo, 1992) or an optimization algorithm based on the foraging behaviour of the swarm of honey bees *artificial bee colony* (Karaboga, 2005) – for example, it could also be successfully engaged to provide robust solutions with more general tasks such as complex problem decomposition based on the principle of flexible specialised labour allocation in the colony or hive. The captivating notion here is of course the reliance on the set of relatively simple operational rules that govern the local behaviour of individual stand-alone agents, successful application of which would result in emergent complex collective intelligence behaviour.

Consequently, at least to some extent for artificial systems, the key problem revolves around determining and accurately defining these sets of rules to facilitate this self-organised global intelligence behaviour, which of course naturally emerged and was refined in biological systems over millennia in the process of natural evolution. With complex systems, it may be extremely difficult to predict and subsequently assess the modelled outcome of final global emergent behaviour based on the set of simple rules – especially if it is an adaptive dynamic system designed to respond quickly to undefined environmental input factors. A number of key principles to keep in mind while considering a swarm intelligence system are as follows: first, very simple rules are able to generate unpredictable and often counterintuitive collective behaviours; second, seemingly minor modifications to these simple sets of rules can result in radically altered collective behaviours as a result; and third, even though the complex task of predicting collective behaviour that may very well be beyond the extent of human intelligence capacity, it is nevertheless possible to observe and predict the modelled collective behaviour in simulated artificial systems. These artificial systems can prove to be invaluable tools to advance and improve our understanding of collective behaviour, and ultimately help predict what collective behaviours would emerge within a certain set of constraints. In an organisational setting, these simple rules may be modelled to predict how simple sets of rules can affect staff knowledge acquisition rates, productivity, loyalty, personnel turnover, and so on – many of which would also apply to consumer environemnt. Indeed, as the global organisations continue to evolve following the technological and communication advancements and ever increasing reliance on the user-provided content and input in the product design and innovation processes, the very notion what constitutes an organisation and its ways of working may be redefined in the future – and swarm intelligence methods could play an imperative role to establish the decentralised self-organising governance rules for such organisations. Another essential consideration when assessing the efficiency and effectiveness of particular rules in a swarm intelligence system – biological or artificial – is the activation mechanism employed to communicate and transfer information between the stand-alone agents or sub-systems, particularly in relation to the overall aim and desired outcome that the emergent swarm intelligence should strive to develop and maximise.

Nevertheless, there are numerous obstacles to adopting swarm intelligence methods – it may be difficult for some to understand the mechanisms and structure of swarm intelligence who are unfamiliar with self-organising systems, the emergent nature of collective behaviour could be increasingly complex to describe and predict, or drawing parallels between groups of individuals such as human consumers and the swarms of insects may not be an appealing concept from the social point of view. Nevertheless in certain circumstantial environments, the collective human behaviour is constrained in a similar manner to the biological swarm systems, and parallels can be drawn to illustrate not only the conceptual, but also practical implications that may be useful to optimise the process and behaviour to achieve higher level of organisation and strategic governance. The swarm intelligence methods discussed thus far would of course be largely applicable when it comes to the discussion of consumer behaviour in general as well, and could be useful to optimise and facilitate and cultivate certain types of desirable behaviours – be that to either benefit the consumer, or maximise the returns of the company, or in an optimal situation provide a substantial benefit to all parties involved, perhaps as a result of optimisation programme for example. The way these algorithms and general approaches could potentially be employed with the research project discussed here to extend the work undertaken thus far is to develop a multitude of stand-alone connectionist networks to represent individual consumers and their purchasing behaviour across all available categories – the dataset employed here utilised but a single product category from a larger data that covers a wider set of consumption behaviours across a large number of product categories which would make it possible to broaden the scope substantially compared to a rather focused single category prototypical consumer approach employed here. These stand-alone connectionist networks could be set to interact between each other within the environment that can be described employing the purchasing context variables. This would arguably represent an artificial system that can better describe the real consumer setting and the purchasing decision-making process within the context of the purchasing setting that would serve as constraints for the stand-alone agent interactions, and can be employed for both predictive and explanatory purposes. Allowing stand-alone agents that represent individual consumers to interact within a consumer setting could illustrate how consumers can influence each other in the process – something that is often ignored or overlooked in marketing research where consumers are taken to operate in a vacuum devoid of any consumer interaction and competition variables. One of the reasons for that is the complexity that is concerned with attempts to provide an adequate account for the continuity of behaviour and the learning history as discussed in detail in the previous chapters.

8.4 Consumer provided content and design

The constantly evolving technological advancements facilitate the emergence of truly global economies and organisations by continuously removing the obstacles that traditionally and historically hindered this progression. Digital media was able to unlock previously unattainable potential to bring together functionally distinct areas of interest and expertise to facilitate the development of collective cognitive problem-solving faculties.

If *collective* human behaviour can be characterised as a set of persons interacting among each other for prolonged periods of time (for an extended discussion on collective behaviour please refer to Krause & Ruxton, 2002), organisations and consumer groups would constitute a prime example of a setting suitable for the emergence of swarm intelligence as it provides an opportunity to solve large scale problems through the facilitation of collective cognitive problem-solving abilities which quite possibly would otherwise be unattainable to individual contributors. As opposed to simpler largely organisms such as ants or bees that for the most part display uniform levels of performance however, there is not only a high level of inter-individual variability among human performance – there are also those individuals who are able to display superior levels of performance overall across multiple measurement criteria. Therefore, large body of psychological research on collective decision-making revolved around the concept of comparative assessment: namely, studying whether collective decision-making would be able to provide adequate or better results than individuals with superior decision-making abilities (as first demonstrated empirically by Galton, 1907). Because of this, swarm intelligence at times could be positioned as an alternative and even a threat to expert centres – indeed, one potential concern is that it can contribute to decomposition and erosion of the expert centres, as diverse groups of individuals that do not normally possess expert knowledge are able to show comparable level of performance with the level exhibited by the experts. This is however incorrect, as swarm intelligence is more appropriate for a particular type of tasks that require a large number of uncorrelated imprecise estimates that eventually result in a close approximation of actual value, whereas the tasks with a large systematic bias that prevent extraction of usable information to solve the task are better suitable for the centres of expertise. As most tasks in an organisational or any other setting would incorporate a degree of both bias and imprecision, it is imperative to identify and select the tasks that are appropriate for swarm intelligence methods, and some qualitative characteristics are important to keep in mind in the process (Krause, Ruxton, & Krause, 2010).

It is often the case that swarm intelligence task groups follow the composition principle based on functional, methodological, and epistemological diversity in attempt to maximise the potential performance. It should be obvious how this potentially could considerably improve some types of performance: for example, programmes on information gathering from broadly diverse individual contributors facilitated by technological advances would enable accessing and processing collective knowledge and learning histories on unprecedented level. Consumer provided content and design are the two areas where organisations gradually adopt swarm intelligence methods to generate insightful and actionable information by making the most of what brand communities are able to offer (Fournier & Lee, 2009). Numerous interactive open-access web-based platforms that became available recently to enable any type of a discussion and problem-solving where vast numbers of individuals are able to contribute in a process that evolves into a partially self-organised and decentralised system. A good example would be the ongoing process of producing and constantly improving and expanding the open-source statistical programing language and programming environment *R,* which is extensively employed here as a method of choice for all statistical modelling, could be viewed as one of the great example of swarm intelligence methods where a type of self-organised and partially decentralised system emerges as a result of a continuous deliberate collective effort from many stand-alone individual contributors.

Innovation is another set of such tasks that normally deal with high levels of imprecision, and could be particularly suitable for the swarm intelligence methods as discussed in the following paragraphs.

8.5 Collective consumer innovation

Recent advances in computer technologies make it seemingly effortless to access global populations, and facilitate information gathering by enabling the process to engage with collective knowledge and creativity of enormous number of individuals in a dynamic and interactive manner that allows collective process of simultaneous collaboration. Collective consumer decision-making that became more evident lately as a result of changing social environment is increasingly accepted as a driving force behind some of the companies from the digital age who recognise that collective involvement of consumer communities could be treated as a novel form of a low-cost resource (McConnell & Huba, 2007). Indeed with the recent upsurge in a number of various open-source projects, progressively high numbers of consumers who are involved in product development and feature modification are being recognised as an imperative part of a collective creative process (Von Hippel, 2005). As an added benefit, this dynamic collaborative creative environment contributes to the traditional word of mouth marketing, while being facilitated and amplified by sophisticated web-based solutions. The distinctive borders between the process of consumption and that of creation and production are becoming increasingly blurred, as the notion of collective consumer innovation allows the immersed organisation of consumer purchasing decision-making and creative innovation to emerge as a combination of the two.

Collective consumer innovation refers to the process that is said to occur when consumers discover new interpretations in the collaborative process of social interaction within a consumer group – something that would be impossible to achieve while thinking on their own: the varying range of backgrounds and experiences should offer increased probability to identify ideas fit to resolve a particular consumption-related task, whereas the accumulating collective experience and knowledge would help establish potential solution selection criteria and mechanisms to develop and realise the idea to be subsequently propagated and promoted to wider consumer populations utilising collective network (Hargadon & Bechky, 2006). For organisations, the appropriate approach may be to position themselves as a part of the creative community rather than attempt to manage the collective creative process – it requires appropriate technological framework to enable the fostering of social and cultural fabric for the collective consumer innovation community (Kozinets, Hemetsberger, & Schau, 2008). Some communities occupy ethical and sustainable consumption viewpoints – something that could perhaps be accounted for as an additional set of constraints or selection criteria as attempted to be modelled using swarm intelligence methods or connectionist network.

8.6 Commercial application and data mining

It would be important to discuss the commercial application of course, as applied deployment in the industry would not only be an excellent practical test to validate the model output with previously unseen real-life consumer data, but would also expand the applied dimension that has the potential to generate innovative insight and suggest novel lines of inquiry for any future research. Moreover, some of the valuable potential data sources compiled and maintained by certain industries highly reliant on data may otherwise be inaccessible. A few potential examples will be discussed very briefly in the following paragraphs that rely heavily on efficient and effective information collection and organisation that could benefit substantially from the integrated connectionist network solutions.

The obvious candidate for connectionist modelling of consumer behaviour is of course the financial industry that historically deals with collecting large amounts of data and information to be used in risk modelling and similar purposes. Big data enhanced by the connectionist modelling could potentially carry huge benefits, financial and otherwise: for example, each regular bankcard payment could be enhanced with a cluster of data to describe in detail a purchased items list containing full product attributes, quantities purchases, trade channel used, and purchasing environment details. This would enable to offer a full account of purchases to consumers to improve the understanding of their spend, improve the unauthorised and fraudulent use tracking and prevention, qualitatively improve the account history tracking and define relationship with the bank, improve the accuracy of forecasting the individual financial performance and behaviour and likely use of any financial services in the future, simplify the credit decision process, and so on.

New customer acquisition and current customer retention with attrition minimisation are the two major ongoing strategic marketing programmes that any marketing focused company should devote a significant amount of resources and effort to manage. Predictive analytics and forecasting modelling are commonly the areas of focused effort where the capacity of connectionist modelling and big data analytics could substantially improve the consumer targeting efforts.

The technological advancements consistently provide cheaper and more efficient methods to collect and store information, and the amount of data being created has been growing exponentially for a number of decades now. Integration of *R* functionality into large-scale applications such as the latest release of *de facto* data industry standard SQL database management software represents a considerable advancement towards improving and developing the data mining capacity. Integrated big data analytics solutions could revolutionise the way the information is collected and processed on a large scale to produce actionable insights useful to inform the process of strategic decision-making.

Any human activity inevitably produces waste – managing and minimising the waste could provide not only the benefit of smaller ecological footprint in market conditions, which are to become increasingly more regulated and therefore costly in the future, but also the immediate financial benefits of improved return on investment and cash flow management. One obvious example would be direct mail acquisition campaigns that produce printed materials, which are then delivered by post. These campaigns are extremely wasteful, where response rates as low as only 1% are to be expected, while the remaining 99% of all printed and delivered materials are normally disposed of as waste. Improving the predictive analytics to allow better consumer targeting would not only reduce the generated printed materials being systematically wasted, but will also reduce the marketing costs.

Classic marketing and consumer behaviour disciplines historically have been largely focused on the purchasing stage of the total consumption process – as is obviously the case in this research project as well. This is of course the case for a number of reasons, one of which is that the consumer purchasing decision is effectively complete when the contract between the buyer and the seller is made, and the money is paid in exchange for a product ownership. The purchasing stage however does not provide a comprehensive account of environmental and social aspects of a total consumption process, and increasingly large numbers of organisations strive to develop measurement criteria and methods to assess the sustainability of their customer base as a strategic long-term sustainable growth solution, where integrated data would be particularly useful.

8.7 Summary

Even though some of the methodologically innovative approaches discussed in this chapter like swarm intelligence could be difficult to implement and challenging to adopt, the potential benefits outweigh those few concerns that are voiced – for example the role of expert centres would evolve rather than diminish and seize to exist towards the type of expertise that enables to facilitate and cultivate the benefits in decision-making that a swarm intelligence method could offer. Different levels to structure and carry out consumer behaviour analysis were also briefly discussed in this chapter, concluding with the overview of possible commercial applications.

9. Concluding remarks

Although this research project at times strived to venture into grander philosophical concepts such as intelligence, cognition, and artificial intelligence, the main scope of the project was intentionally constrained to the boundaries of the field of consumer behaviour – boundaries which themselves could be rather blurry at times and generally welcome participation of a number of disciplines: psychology, marketing, economics, artificial intelligence as the case here, and many others depending on the particular application. In these final pages, it would be worthwhile to revisit these broader issues and reiterate particular points made throughout the work presented here.

9.1 Contributions

One of the first points that this research project examined was the comparative examination that assessed how connectionist neural network models measure up against traditionally employed methods of analysis such as logistic regression. Not only predictive, but also explanatory dimension was of interest, where it was shown that simple connectionist models that do not incorporate hidden layers provide connection weights that are analogous to the coefficient values of the regression, suggesting analogous level of performance that the two methods are able to provide when it comes to predictive capacity. When it comes to the explanatory capacity however, the regression is already in its final form and does not provide any means for further developed – whereas connectionist models is only in its most primitive form, and can be developed substantially by incorporating hidden layers and growing the network.

Second point is the fact that this research project was able to directly contribute to development and advancement of a number of statistical programming packages in *R*, namely *RSNNS* and *NeuralNetTools*, which are now available for all researchers to use and potentially contribute to the understanding of connectionist frameworks going forward.

Third, a number of advanced connectionist models of consumer behaviour have been developed, ranging in size and complexity of the architecture, which provide empirical evidence to corroborate previous research findings and help explain and guide further research.

Fourth, a simulated dataset was developed to assess the efficiency and accuracy of the pruning algorithms that was able to show in an obvious manner and provide convincing empirical evidence for the effectiveness of pruning algorithms in optimising the network architecture to expose the core underlying architecture in attempt to improve the exploratory and interpretative function of a connectionist model.

Fifth, once it was evident that pruning algorithms work very well and as designed, the varying number of iterations and retrain cycles were examined in attempt to investigate the level of influence these parameters are able to exert over the overall connectionist model capacity to predict and explain consumer behaviour, which could be further explored in a string of research that would focus on network architecture optimisation to minimise the level of recourses required to perform these advanced statistical models.

Sixth, the initial network architecture designs were examined to explore and compare the subsequent network learning process and final form. This could be a useful line of research to extend in attempt to identify the optimal learning methods and initial network architecture design for a particular problematic.

Seventh, pruning algorithms were employed to optimise the network architecture for explanatory and interpretative purposes – it is argued here that this offers an alternative method to a variable contribution analysis, and could be a particularly useful technique to explore complex phenomena such as consumer decision-making as an emergent process.

Eighth, research carried out here provides empirical evidence to support the proposition that connectionism is a robust and coherent approach that is particularly suited to extend the theoretical framework of BPM.

Ninth, connectionist models developed here provide empirical evidence that informational and utilitarian reinforcement can be observed as an emergent process by means of distributed representation – an inherent functionality of connectionist networks to embody complex phenomena.

Tenth, as part of this research programme, it reinforced a tendency towards addressing the notion of continuity of behaviour from the radical behaviourism point of view with the adoption of intentional behaviourism.

Eleventh, the theoretical and methodological deliberations contemplated here propose a general structure to promote the move towards addressing the notion of learning history as part of overall consumer behaviour process – naturally there is a large amount of future work required here.

Twelfth, the detailed overview of the field of artificial intelligence in relation to the process of consumer behaviour is offered here, in attempt to encourage the future symbiotic collaborative efforts that could advance the developments in both fields of study.

9.2 Summary

In the first chapter, the research project was briefly introduced, outlining the contents and structure. The overall motivation for the research project was presented, offering the scope and the summary of the work to be carried out. Next two chapters focused on the fields of consumer behaviour and artificial intelligence, and offered a comprehensive review that also touched upon a number of related disciplines in the course of inquiry. Multidisciplinary nature of work described here embraced the philosophical aspects of critical behaviourism and cognitive sciences, considered and reviewed modelling approached that follow both traditional symbolic and connectionist neural networks designs, and contemplated the theoretical frameworks that propose to extend the theory of Behavioural Perspective Model into the realm of connectionist architectures. For that reason, an extended discussion that overviews the field of consumer behaviour was offered, followed by an overview of theoretical and philosophical frameworks of radical behaviourism, and continued to describe the details of Behavioural Perspective Model as a next step towards an interpretative model of consumer behaviour. The discussion was followed by an extensive discussion of the science of the artificial, introducing the field of artificial intelligence: the predictive and explanatory capacity of symbolic modelling methods was discussed at length and compared against the connectionist neural networks approach, describing in detail the techniques and architecture optimisation algorithms employed in neural network models. Chapter 4 focused on the research methods and outlined in detail the research questions, the research methods employed in the course of the research project, and the philosophical position adopted here. This was followed by sections that described the methods, dataset structures and variables, modelling techniques, and the sequence of the research process. Once the research methods were clearly explained, the next chapter described the modelling methods that were carried out here, and offered a detailed account of the statistical analyses developed throughout this research project. Specific testing procedures to advance the line of inquiry were explained in detail, and offered an overview of the results. Chapter 6 discussed the findings, and offered an interpretative account of results within the wider context of consumer behaviour. Variable contribution analysis as a method to improve the descriptive functions of the modelling was discussed in light of employing the advanced connectionist modelling method, posing an argument that connectionist neural networks approach is particularly appropriate to provide the comprehensive explanatory and interpretative account of consumer behaviour, where pruning algorithms were employed to optimise the network architecture to expose the core architecture. This then was followed by a discussion around the theoretical implications. Critical assessment of the research project was offered in the following chapter to demonstrate precision, thoroughness, and level of contribution in comparison with its closest rival, the tradition of cognitive science. Chapter 8 was developed around the possible future research directions which were briefly discussed, identifying a number of possible strings of inquiry that ranged from commercial in nature that aimed to apply and test the methods proposed here in the applied industry setting, to typically theoretical and philosophical endeavours that aimed to explore the concept of distributed representation further and propose to potentially extend the line of inquiry into the field of swarm intelligence. The final chapter offered closing remarks, revisiting the contributions this research project aims to offer, and concluding with an overall summary of the project.

References

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines*. *Cognitive Science, 9*(1), 147-169.

Adya, M., & Collopy, F. (1998). How effective are neural networks at forecasting and prediction? A review and evaluation. *J. Forecasting, 17*, 481-495.

Ainslie, A., & Rossi, P. E. (1998). Similarities in choice behavior across product categories. *Marketing Science, 17*(2), 91-106.

Alvesson, M., & Deetz, S. (2000). *Doing critical management research*. London: Sage.

Ametller, L., Garrido, L., Stimpfl-Abele, G., Talavera, P., & Yepes, P. (1996). Discriminating signal from background using neural networks: Application to top-quark search at the Fermilab Tevatron. *Physical Review D, 54*(1), 1233.

Anderson, J. A., & Hinton, G. E. (1981). Models of information processing in the brain. *Parallel models of associative memory*, 9-48.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive psychology, 6*(4), 451-474.

Anderson, J. R. (1981). *Cognitive skills and their acquisition*: Psychology Press.

Anderson, J. R. (1983a). The architecture of cognition. *Cambridge, Mass.: HalYard University Pres*.

Anderson, J. R. (1983b). A spreading activation theory of memory. *Journal of verbal learning and verbal behavior, 22*(3), 261-295.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological review, 94*(2), 192.

Anderson, J. R., & Pirolli, P. L. (1984). Spread of activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 791.

Anderson, P. F. (1983). Marketing, scientific progress, and scientific method. *The Journal of Marketing, 47*(4), 18-31.

Anderson, P. F. (1986a). On method in consumer research: a critical relativist perspective. *Journal of Consumer Research, 13*(2), 155-173.

Anderson, P. F. (1986b). On method in consumer research: a critical relativist perspective. *The Journal of Consumer Research, 13*(2), 155-173.

Anderson, P. F. (1988a). Relative to what - that is the question: a reply to Siegal. *Journal of Consumer Research, 15*(1), 133-137.

Anderson, P. F. (1988b). Relativism revidivus: in defense of critical relativism. *Journal of Consumer Research, 15*(3), 403-406.

Andersson, F. O., Åberg, M., & Jacobsson, S. P. (2000). Algorithmic approaches for studies of variable influence, contribution and selection in neural networks. *Chemometrics and intelligent laboratory systems, 51*(1), 61-72.

Aoki, I., & Komatsu, T. (1997). Analysis and prediction of the fluctuation of sardine abundance using a neural network. *Oceanolica Acta, 20*(1), 81-88.

Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition, 13*(3), 263-308.

Barsalou, L. W. (1983). Ad hoe categories. *Memory & cognition, 11*(3), 211-227.

Bashford, S. (2009). Bring me sunshine. *Marketing*, 32-33.

Bates, E., Bretherton, I., & Snyder, L. (1988). From first words to grammar: Cambridge: Cambridge University Press.

Beck, M. (2015). NeuralNetTools: Visualization and Analysis Tools for Neural Networks.

Beni, G., & Wang, J. (1993). Swarm intelligence in cellular robotic systems *Robots and Biological Systems: Towards a New Bionics?* (pp. 703-712): Springer.

Bergmeir, C., & Benítez, J. M. (2012). Neural networks in R using the Stuttgart neural network simulator: RSNNS. *Journal of Statistical Software, 46*(7), 1-26.

Bernard, H. R. (1995). *Research methods in anthropology: Qualitative and quantitative approaches*. Walnut Creek, CA: AltaMira Press.

Bishop, C. M. (1995). *Neural networks for pattern recognition*: Oxford university press.

Bloom, L. (1970). Language development: Form and function in emerging grammars. MIT Research Monograph, No 59.

Bryman, A., & Bell, E. (2007). *Business research methods*. Oxford; New York: Oxford University Press.

Bybee, J. L., & Slobin, D. I. (1982). Rules and schemas in the development and use of the English past tense. *Language*, 265-289.

Calder, B. J., & Tybout, A. M. (1987). What Consumer Research Is. *Journal of Consumer Research, 14*(1), 136-140.

Chen, D., & Ware, D. (1999). A neural network model for forecasting fish stock recruitment. *Canadian Journal of Fisheries and Aquatic Sciences, 56*(12), 2385-2396.

Chib, S., Seetharaman, P., & Strijnev, A. (2002). Analysis of multi-category purchase incidence decisions using IRI market basket data. *Advances in Econometrics, 16*, 57-92.

Chintagunta, P. K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science, 12*(2), 184-208.

Chomsky, N. (1957). Syntactic Structures. *The Netherlands*.

Chomsky, N. (1959). A review of BF Skinner's Verbal Behavior. *Language, 35*(1), 26-58.

Chomsky, N. (1968). Language and mind. *Psychology Today, 1*(9), 48-&.

Chung, J., & Rao, V. R. (2003). A general choice model for bundles with multiple-category products: Application to market segmentation and optimal pricing for bundles. *Journal of Marketing Research, 40*(2), 115-130.

Cibas, T., Soulié, F. F., Gallinari, P., & Raudys, S. (1996). Variable selection with neural networks. *Neurocomputing, 12*(2–3), 223-248.

Cleeremans, A., Servan-Schreiber, D., & McClelland, J. L. (1989). Finite state automata and simple recurrent networks. *Neural computation, 1*(3), 372-381.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological review, 100*(4), 589.

Cornwell, B., Chi Cui, C., Mitchell, V., Schlegelmilch, B., Dzulkiflee, A., & Chan, J. (2005). A cross-cultural study of the role of religion in consumers' ethical positions. *International Marketing Review, 22*(5), 531-546. doi:10.1108/02651330510624372

Crosas, M., King, G., Honaker, J., & Sweeney, L. (2015). Automating Open Science for Big Data. *The ANNALS of the American Academy of Political and Social Science, 659*(1), 260-273.

Cunningham, L. F., Young, C. E., Moonkyu, L., & Ulaga, W. (2006). Customer perceptions of service dimensions: cross-cultural analysis and perspective. *International Marketing Review, 23*(2), 192-210. doi:10.1108/02651330610660083

Curry, B., Davies, F., Phillips, P., Evans, M., & Moutinho, L. (2001). The Kohonen self-organizing map: an application to the study of strategic groups in the UK hotel industry. *Expert systems, 18*(1), 19-31.

Curry, B., & Moutinho, L. (1993). Neural Networks in Marketing: Modelling Consumer Responses to Advertising Stimuli. *European Journal of Marketing, 27*(7), 5-20. doi:10.1108/03090569310040325

Dennett, D. C. (1981). *Brainstorms: Philosophical essays on mind and psychology*: MIT press.

Dennett, D. C. (2002). *Content and consciousness*: Routledge.

Denzin, N., & Lincoln, Y. (2000). The discipline and practice of qualitative research. *Handbook of qualitative research, 2*, 1-28.

Despagne, F., & Massart, D.-L. (1998). Variable selection for neural networks in multivariate calibration. *Chemometrics and intelligent laboratory systems, 40*(2), 145-163.

Dimopoulos, I., Chronopoulos, J., Chronopoulou-Sereli, A., & Lek, S. (1999). Neural network models to study relationships between lead concentration in grasses and permanent urban descriptors in Athens city (Greece). *Ecological modelling, 120*(2), 157-165.

Dorigo, M. (1992). Optimization, learning and natural algorithms. *Ph. D. Thesis, Politecnico di Milano, Italy*.

Dreyfus, H. (1992). *What computers still can't do: a critique of artificial reason*: MIT press.

Dreyfus, H., Dreyfus, S. E., & Athanasiou, T. (2000). *Mind over machine*: Simon and Schuster.

Durkheim, E. (1964). *The division of labor in society*. New York: Free Press.

East, R., Hammond, K., & Wright, M. (2007). The relative incidence of positive and negative word of mouth: A multi-category study. *International journal of research in marketing, 24*(2), 175-184.

Elman, J. L. (1989). *Structured representations and connectionist models*. Retrieved from

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*(2), 179-211.

Erdem, T. (1998). An empirical analysis of umbrella branding. *Journal of Marketing Research*, 339-351.

Feyerabend, P. (1975). *Against method: Outline of an anarchistic theory of knowledge*. London: Verso.

Feyerabend, P. (1978). *Science in a free society*. London: NLB.

Feyerabend, P. (1987). *Farewell to reason*. London: Verso.

Fodor, J. A. (1975). *The language of thought* (Vol. 5): Harvard University Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*(1), 3-71.

Foucault, M. (1995). *Discipline and punish: the birth of the prison*. New York: Vintage Books.

Fournier, S., & Lee, L. (2009). Getting brand communities right. *Harvard business review, 87*(4), 105-111.

Foxall, G. R. (1990). *Consumer psychology in behavioral perspective*: Beard Books.

Foxall, G. R. (2004). *Context and cognition: Interpreting complex behavior*: New Harbinger Publications.

Foxall, G. R. (2005). *Understanding consumer choice*: Palgrave Macmillan.

Foxall, G. R. (2009). *Interpreting consumer choice: The behavioural perspective model*: Routledge.

Foxall, G. R. (2016). *Addiction as Consumer Choice: Exploring the Cognitive Dimension*: Routledge.

Foxall, G. R., Wells, V. K., Chang, S. W., & Oliveira-Castro, J. M. (2010). Substitutability and independence: Matching analyses of brands and products. *Journal of Organizational Behavior Management, 30*(2), 145-160.

Foxall, G. R., Yan, J., Oliveira-Castro, J. M., & Wells, V. K. (2013). Brand-related and situational influences on demand elasticity. *Journal of Business Research, 66*(1), 73-81.

Fullerton, R. A. (1987). Historicism: What it is, and what it means for consumer research. *Advances in Consumer Research, 14*(1), 431-434.

Gallant, S. I. (1993). *Neural network learning and expert systems*: MIT press.

Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature, 75*, 450-451.

Garson, D. G. (1991). Interpreting neural network connection weights.

Gevrey, M., Dimopoulos, I., & Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling, 160*(3), 249-264.

Ghauri, P. N., & Grønhaug, K. (2005). *Research methods in business studies: a practical guide*: Prentice Hall.

Goh, A. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering, 9*(3), 143-151.

Goulding, C. (1999). Consumer research, interpretive paradigms and methodological ambiguities. *European Journal of Marketing, 33*(9/10), 859-873.

Greene, M. N. (2011). Behavioural Perspective Model: A connectionist view (Master's thesis). Cardiff University, Cardiff, UK.

Greenland, S. (1989). Modeling and variable selection in epidemiologic analysis. *American journal of public health, 79*(3), 340-349.

Guégan, J.-F., Lek, S., & Oberdorff, T. (1998). Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature, 391*(6665), 382-384.

Güneren, E., & Öztüren, A. (2008). Influence of Ethnocentric Tendency of Consumers on Their Purchase Intentions in North Cyprus. *Journal of Euromarketing, 17*(3/4), 219-231.

Haas, J. (1987). Prospects for Consumer Research. *Advances in Consumer Research, 14*, 76-78.

Haigh, J., & Crowther, G. (2005). Interpreting motorcycling through its embodiment in life story narratives. *Journal of Marketing Management, 21*(5/6), 555-572.

Hanson, N. R. (1958). *Patterns of Discovery, an Inquiry Into the Conceptual Foundations of Science, by Norwood Russell Hanson*: University Press.

Hargadon, A. B., & Bechky, B. A. (2006). When collections of creatives become creative collectives: A field study of problem solving at work. *Organization Science, 17*(4), 484-500.

Harris, A. J., & Hahn, U. (2011). Unrealistic optimism about future life events: a cautionary note. *Psychological review, 118*(1), 135.

Hassibi, B., & Stork, D. G. (1993). *Second order derivatives for network pruning: Optimal brain surgeon*: Morgan Kaufmann.

Hassibi, B., Stork, D. G., Wolff, G., & Watanabe, T. (1994). Optimal Brain Surgeon: Extensions and performance comparison.

Hassibi, B., Stork, D. G., & Wolff, G. J. (1993). *Optimal brain surgeon and general network pruning.* Paper presented at the Neural Networks, 1993., IEEE International Conference on.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological approach*: John Wiley & Sons.

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*: Psychology Press.

Heidegger, M. (1988). *The basic problems of phenomenology* (Vol. 478): Indiana University Press.

Heravi, S., & Morgan, P. (2014a). A comparison of six sampling schemes for price index construction in a COICOP food group. *Applied Economics, 46*(22), 2685-2699.

Heravi, S., & Morgan, P. (2014b). Sampling schemes for price index construction: a performance comparison across the classification of individual consumption by purpose food groups. *Journal of Applied Statistics, 41*(7), 1453-1470.

Hildum, D. C., & Brown, R. W. (1956). Verbal reinforcement and interviewer bias. *Journal of Abnormal and Social Psychology, 53*, 108-111.

Hinton, G. E. (1986). *Learning distributed representations of concepts.* Paper presented at the Proceedings of the eighth annual conference of the cognitive science society.

Hinton, G. E., Sejnowski, T. J., & Ackley, D. H. (1984). *Boltzmann machines: Constraint satisfaction networks that learn*: Carnegie-Mellon University, Department of Computer Science Pittsburgh, PA.

Holbrook, M. B. (1987). What is consumer research? *Journal of Consumer Research, 14*(1), 128-132.

Holbrook, M. B. (1989). Seven routes to facilitating the semiological interpretation of consumption symbolism and marketing imagery in works of art: Some tips for wildcats. *Advances in Consumer Research, 16*(1), 420-425.

Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *The Journal of Consumer Research, 9*(2), 132-140.

Holbrook, M. B., & O'Shaughnessy, J. (1988). On the scientific status of consumer research and the need for an interpretive approach to studying consumption behavior. *Journal of Consumer Research, 15*(3), 398-402.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). Induction: Cambridge, MA: MIT Press.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences, 79*(8), 2554-2558.

Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: a market comparative study. *Decision support systems, 37*(4), 543-558.

Hudson, L. A., & Ozanne, J. L. (1988). Alternative ways of seeking knowledge in consumer research. *Journal of Consumer Research, 14*(4), 508-521.

Hunt, S. D. (1990). Truth in marketing theory and research. *Journal of Marketing, 54*(3), 1.

Hunt, S. D. (1991). Positivism and paradigm dominance in consumer research: Toward critical pluralism and rapprochement. *Journal of Consumer Research, 18*(1), 32-44.

Hunt, S. D. (1993). Objectivity in marketing theory and research. *Journal of Marketing, 57*(2), 76.

Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy*: Northwestern University Press.

Husserl, E. (2012). *Ideas: General introduction to pure phenomenology*: Routledge.

Huttenlocher, J., Higgins, E. T., & Clark, H. H. (1971). Adjectives, comparatives, and syllogisms. *Psychological review, 78*(6), 487.

Insko, C. A. (1965). Verbal reinforcement of attitude. *Journal of Personality and Social Psychology, 2*, 621-623.

Iyengar, R., Ansari, A., & Gupta, S. (2003). Leveraging information across categories. *Quantitative Marketing and Economics, 1*(4), 425-465.

Karaboga, D. (2005). *An idea based on honey bee swarm for numerical optimization*. Retrieved from

Kaynak, E., & Kara, A. (2001). An examination of the relationship among consumer lifestyles, ethnocentrism, knowledge structures, attitudes and behavioural tendencies: a comparative study in two CIS states. *International Journal of Advertising, 20*(4), 455-482.

Kehret-Ward, T. (1988). Using a semiotic approach to study the consumption of functionally related products. *International Journal of Research in Marketing, 4*(3), 187-200.

Kerlinger, F. N. (1964). *Foundations of behavioral research: Educational and psychological inquiry*. New York; London: Holt, Rinehart and Winston.

Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE, 78*(9), 1464-1480.

Kohonen, T. (1998). The self-organizing map. *Neurocomputing, 21*(1), 1-6.

Kosslyn, S. M. (1980). *Image and mind*: Harvard University Press.

Kozinets, R. V., Hemetsberger, A., & Schau, H. J. (2008). The wisdom of consumer crowds collective innovation in the age of networked marketing. *Journal of Macromarketing, 28*(4), 339-354.

Krause, J., & Ruxton, G. D. (2002). *Living in groups*: Oxford University Press.

Krause, J., Ruxton, G. D., & Krause, S. (2010). Swarm intelligence in animals and humans. *Trends in ecology & evolution, 25*(1), 28-34.

Kuczaj, S. A. (1977). The acquisition of regular and irregular past tense forms. *Journal of verbal learning and verbal behavior, 16*(5), 589-600.

Kuhn, T. S. (1996). *The structure of scientific revolutions*. Chicago ; London: University of Chicago Press.

Kuhn, T. S. (2012). *The structure of scientific revolutions*: University of Chicago press.

Kumcu, E. (1987). An historical perspective framework to study consumer behavior and retailing systems. *Advances in Consumer Research, 14*(1), 439-441.

LeCun, Y., Denker, J. S., Solla, S. A., Howard, R. E., & Jackel, L. D. (1989). *Optimal brain damage.* Paper presented at the NIPs.

Lek, S., Belaud, A., Baran, P., Dimopoulos, I., & Delacoste, M. (1996). Role of some environmental variables in trout abundance models using neural networks. *Aquatic Living Resources, 9*(01), 23-29.

Lek, S., Belaud, A., Dimopoulos, I., Lauga, J., & Moreau, J. (1995). Improved estimation, using neural networks, of the food consumption of fish populations. *Marine and Freshwater Research, 46*(8), 1229-1236.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., & Aulagnier, S. (1996). Application of neural networks to modelling nonlinear relationships in ecology. *Ecological modelling, 90*(1), 39-52.

Locke, J. (1997). *An essay concerning human understanding*. London ; New York: Penguin.

Long, A. (1998). Plato's Apologies and Socrates in the Theaetetus. *Method in Ancient Philosophy*, 113–136.

Lu Hsu, J., & Han-Peng, N. (2008). Who are ethnocentric? Examining consumer ethnocentrism in Chinese societies. *Journal of Consumer Behaviour, 7*(6), 436-447. doi:10.1002/cb.262

Luger, G. F. (2005). *Artificial intelligence: structures and strategies for complex problem solving*: Pearson education.

Lutz, R. J. (1989). Positivism, naturalism and pluralism in consumer research: Paradigms in paradise. *Advances in Consumer Research, 16*(1), 1-8.

Ma, Y., Seetharaman, S., & Narasimhan, C. (2005). Empirical analysis of competitive pricing strategies with complementary product lines. *Available at SSRN 899606*.

Maier, H. R., & Dandy, G. C. (1996). The use of artificial neural networks for the prediction of water quality parameters. *Water resources research, 32*(4), 1013-1022.

Manchanda, P., Ansari, A., & Gupta, S. (1999). The “shopping basket”: A model for multicategory purchase incidence decisions. *Marketing Science, 18*(2), 95-114.

Margolis, H. (1987). *Patterns, thinking, and cognition: A theory of judgment*: University of Chicago Press.

Marsden, D., & Littler, D. (1996). Evaluating alternative research paradigms: a market-oriented framework. *Journal of Marketing Management, 12*(7), 645-655.

Mastrorillo, S., Lek, S., & Dauba, F. (1997). Predicting the abundance of minnow Phoxinus phoxinus (Cyprinidae) in the River Ariege (France) using artificial neural networks. *Aquatic Living Resources, 10*(03), 169-176.

Mastrorillo, S., Lek, S., Dauba, F., & Belaud, A. (1997). The use of artificial neural networks to predict the presence of small-bodied fish in a river. *Freshwater biology, 38*(2), 237-246.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological review, 86*(4), 287.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review, 88*(5), 375.

McClelland, J. L., Rumelhart, D. E., & Group, P. R. (1986). Parallel distributed processing. *Explorations in the microstructure of cognition, 2*.

McConnell, B., & Huba, J. (2007). *Citizen marketers: When people are the message*: Kaplan Pub.

McKee, A. (1984). Social economy and the theory of consumer behavior. *International Journal of Social Economics, 11*(3/4), 45.

Medin, D. L. (1989). Concepts and conceptual structure. *American psychologist, 44*(12), 1469.

Microsoft-Corporation. (2006). Microsoft Office Excel 2007.

Minsky, M. (1975). A framework for representing knowledge.

Minsky, M., & Papert, S. (1969). Perceptrons MIT press. *Cambridge, Ma*.

Moutinho, L., Davies, F., & Curry, B. (1996). The impact of gender on car buyer satisfaction and loyalty: A neural network analysis. *Journal of Retailing and Consumer Services, 3*(3), 135-144.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the society for research in child development*, 1-135.

Newell, A. (1994). *Unified theories of cognition*: Harvard University Press.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM, 19*(3), 113-126.

Newell, A., & Simon, H. A. (1981). Computer science as empirical inquiry: Symbols and search. *Mind Design*, 4l.

Nola, R. (1988). *Relativism and realism in science*. Dordrecht; Boston: Kluwer Academic Publishers.

Nord, L. I., & Jacobsson, S. P. (1998). A novel method for examination of the variable contribution to computational neural network models. *Chemometrics and intelligent laboratory systems, 44*(1), 153-160.

Norman, D. A., Rumelhart, D. E., & Group, L. R. (1975). *Explorations in cognition*: WH Freeman San Francisco.

O'Shaughnessy, J. (1985). A return to reason in consumer behaviou: an hermeneutical approach. *Advances in Consumer Research, 12*(1), 305-311.

Olden, J. D., & Jackson, D. A. (2001). Fish–habitat relationships in lakes: gaining predictive and explanatory insight by using artificial neural networks. *Transactions of the American Fisheries Society, 130*(5), 878-897.

Olden, J. D., & Jackson, D. A. (2002). Illuminating the" black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling, 154*(1), 135-150.

Olden, J. D., Joy, M. K., & Death, R. G. (2004). An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling, 178*(3-4), 389-397.

Oliveira-Castro, J. M., Foxall, G. R., & Wells, V. K. (2010). Consumer brand choice: Money allocation as a function of brand reinforcing attributes. *Journal of Organizational Behavior Management, 30*(2), 161-175.

Özesmi, S. L., & Özesmi, U. (1999). An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological modelling, 116*(1), 15-31.

Pachauri, M. (2002). Consumer behavior a literature review. *Marketing Review, 2*(3), 319.

Paivio, A. (2013). *Imagery and verbal processes*: Psychology Press.

Peter, J. P. (1992). Realism or relativism for marketing theory and research: a comment on hunt's "Scientific Realism". *Journal of Marketing, 56*(2), 72-79.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*(1), 73-193.

Pitts, W., & McCulloch, W. S. (1947). How we know universals the perception of auditory and visual forms. *The Bulletin of mathematical biophysics, 9*(3), 127-147.

Plunkett, K., & Marchman, V. (1989). Pattern association in a back propagation network: implications for child language acquisition. *Center for Research in Language. Technical report, 8902*.

Popper, K. R. (1959). *The logic of scientific discovery*. New York: Basic Books.

Proriol, J. (1995). Selection of variables for neural network analysis Comparisons of several methods with high energy physics data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 361*(3), 581-585.

Prus, R., & Frisby, W. (1987). Marketplace dynamics: the P's of "people" and "process". *Advances in Consumer Research, 14*(1), 61-65.

Quine, W. V. O., Churchland, P. S., & Follesdal, D. (2013). *Word and object*: MIT press.

R_Core_Team. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012: ISBN 3-900051-07-0.

R_Studio. (2012). R Studio: integrated development environment for R: Version 0.97.

Rachlin, H. (1994). *Behavior and mind: The roots of modern psychology*: Oxford University Press.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of experimental psychology, 81*(2), 275.

Rosch, E., & Lloyd, B. B. (1978). Cognition and categorization. *Hillsdale, New Jersey*.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive psychology, 7*(4), 573-605.

Rosenblatt, F. (1958). *Two theorems of statistical separability in the perceptron*: United States Department of Commerce.

Rumelhart, D. E. (1975). Notes on a schema for stories. *Representation and understanding: Studies in cognitive science, 211*, 236.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation*. Retrieved from

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling, 5*, 3.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review, 89*(1), 60.

Rumelhart, D. E., & McClelland, J. L. (1985a). Levels indeed! A response to Broadbent.

Rumelhart, D. E., & McClelland, J. L. (1985b). *On learning the past tenses of English verbs*. Retrieved from

Rumelhart, D. E., McClelland, J. L., & Group, P. R. (1988). *Parallel distributed processing* (Vol. 1): IEEE.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. (1986). Sequential thought processes in PDP models. *V, 2*, 3-57.

Russell, G. J., & Kamakura, W. A. (1998). Modeling multiple category brand preference with household basket data. *Journal of Retailing, 73*(4), 439-461.

Russell, G. J., & Petersen, A. (2000). Analysis of cross category dependence in market basket selection. *Journal of Retailing, 76*(3), 367-392.

Ryle, G. (2009). *The concept of mind*: Routledge.

Sanders, C. R. (1987). Consuming as social action: Ethnographic methods in consumer research. *Advances in Consumer Research, 14*(1), 71-75.

Saunders, M., Lewis, P., & Thornhill, A. (2009). *Research methods for business students*. Harlow: Pearson.

Scardi, M., & Harding, L. W. (1999). Developing an empirical model of phytoplankton primary production: a neural network case study. *Ecological modelling, 120*(2), 213-223.

Sejnowski, T. J., & Rosenberg, C. R. (1987). Parallel networks that learn to pronounce English text. *Complex systems, 1*(1), 145-168.

Selfridge, O. G. (1958). Pandemonium: a paradigm for learning in mechanisation of thought processes.

Shepard, R. N. (1989). Internal representation of universal regularities: A challenge for connectionism.

Siegel, H. (1988). Relativism for Consumer Research? (Comments on Anderson). *Journal of Consumer Research, 15*(1), 129-132.

Simon, H. A. (1977). The logic of heuristic decision making *Models of discovery* (pp. 154-175): Springer.

Singh, V. P., Hansen, K. T., & Gupta, S. (2005). Modeling preferences for common attributes in multicategory brand choice. *Journal of Marketing Research, 42*(2), 195-209.

Skinner, B. F. (1938). The behavior of organisms: An experimental analysis.

Skinner, B. F. (1953). *Science and human behavior*: Simon and Schuster.

Skinner, B. F. (2014). *Verbal behavior*: BF Skinner Foundation.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology, 28*(1), 1-39. doi:doi:10.1146/annurev.ps.28.020177.000245

Smolensky, P. (1988a). The constituent structure of connectionist mental states: A reply to Fodor and Pylyshyn. *The Southern Journal of Philosophy, 26*(S1), 137-161.

Smolensky, P. (1988b). On the proper treatment of connectionism. *Behavioral and brain sciences, 11*(01), 1-23.

Song, J., Kong, D., & Yu, J. (1988). Population system control. *Mathematical and Computer Modelling, 11*, 11-16.

SPSS-Inc. (2007). *SPSS Statistics Base 17.0 User's Guide*. Chicago, IL.

Taylor, W. K. (1956). Electrical simulation of some nervous system functional activities. *Information theory, 3*, 314-328.

Touretzky, D. S., & Hinton, G. E. (1988). A distributed connectionist production system. *Cognitive Science, 12*(3), 423-466.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology, 5*(2), 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science, 185*(4157), 1124-1131.

Ullman, M., Pinker, S., Hollander, M., Prince, A., & Rosen, T. (1989). *Growth of regular and irregular vocabulary and the onset of overregularization.* Paper presented at the 14th Boston University Conference on Language Development.

van Kenhove, P., Vermeir, I., & Verniers, S. (2001). An Empirical Investigation of the Relationship between Ethical Beliefs, Ethical Ideology, Political Preference and Need for Closure. *Journal of Business Ethics, 32*(4), 347-361.

Venables, W., & Ripley, B. (2002). Modern Applied Statistics Using S: Springer, New York, NY, USA.

Von Hippel, E. A. (2005). Democratizing innovation.

Watson, J. J., & Wright, K. (2000). Consumer ethnocentrism and attitudes toward domestic and foreign products. *European Journal of Marketing, 34*(9/10), 1149.

Wolcott, H. (2002). Writing up qualitative research... better. *Qualitative Health Research, 12*(1), 91.

Yan, J., Foxall, G. R., & Doyle, J. R. (2012a). Patterns of Reinforcement and the Essential Value of Brands: II. Evaluation of a Model of Consumer Choice. *Psychological Record, 62*(3), 377.

Yan, J., Foxall, G. R., & Doyle, J. R. (2012b). Patterns of reinforcement and the essential values of brands: i. incorporation of utilitarian and informational reinforcement into the estimation of demand. *The Psychological Record, 62*(3), 361.

Zell, A., Mache, N., Hübner, R., Mamier, G., Vogt, M., Schmalzl, M., & Herrmann, K.-U. (1994). SNNS (stuttgart neural network simulator) *Neural Network Simulation Environments* (pp. 165-186): Springer.

Zell, A., Mache, N., Sommer, T., & Korb, T. (1991a). *Design of the SNNS neural network simulator.* Paper presented at the 7. Österreichische Artificial-Intelligence-Tagung/Seventh Austrian Conference on Artificial Intelligence.

Zell, A., Mache, N., Sommer, T., & Korb, T. (1991b). *Recent developments of the SNNS neural network simulator.* Paper presented at the Applications of Artificial Neural Networks II.

Zell, A., Mache, N., Sommer, T., & Korb, T. (1991c). The SNNS neural network simulator *GWAI-91 15. Fachtagung für Künstliche Intelligenz* (pp. 254-263): Springer.

Zell, A., Mache, N., Vogt, M., & Hüttel, M. (1993). *Problems of massive parallelism in neural network simulation.* Paper presented at the Neural Networks, 1993., IEEE International Conference on.