Behavioural and neurological measures of nodal distance in equivalence class formation.

Wang, Ting

Swansea University
Prifysgol Abertawe

Behavioural and Neurological Measures of Nodal Distance in Equivalence Class Formation

Ting Wang

Submitted to the University of Wales in fulfillment of the requirements for the Degree of Doctor of Philosophy (PhD)

Swansea University

June 2010

Supervision provided by Dr. Louise McHugh

Summary (Abstract)

Stimulus equivalence is a behavioural phenomenon in which participants trained in a particular relational pattern (e.g., A → B, B→ C) show a series of additional derived relations including reflexivity (e.g., A→ A), symmetry (e.g., B → A) and transitivity (e.g., A → C). Stimulus equivalence provides a behavioural model of semantic network growth, that is, how new words might enter a network. The current thesis aimed to draw on evidence from the behavioural, cognitive, and neurological literature to provide a coherent account of semantic network growth. There has been much speculation in the stimulus equivalence literature regarding the factors that affect equivalence class formation across both methodological and conceptual levels. Chapters 2 and 3 of the current thesis aimed to single out nodal number effects from methodological confounds, specifically, the training structure and unequal reinforcement history during conditional discrimination training (Experiment 2.1 to 2.4) and unequal stimulus presentation during conditional discrimination training (Experiment 3.1). The results suggested that the nodal number effect is either not a function of differential reinforcement or differential discrimination during conditional discrimination training. Chapters 4 and 5 aimed to integrate equivalence studies with mainstream cognitive and neurological procedures in an attempt to bridge the gap between the semantic network and equivalence literatures. Experiment 4.1 (Chapter 4) investigated the relative contribution and interaction between the factors of number of reinforced trials, time of acquisition, and number of stimulus presentations in equivalence formation / semantic network growth. The results from this work suggest that the number of reinforcers delivered is the most important factor when compared to age of acquisition, and word frequency. Experiment 5.1 and 5.2 (Chapter 5) extended the behavioural literature on equivalence class formation with the addition of a neurological measure, in which, reinforcement and number of trial presentation were manipulated during baseline conditional discrimination training (Experiment 5.1). Participants' response times (RTs) demonstrated priming effects except when participants were exposed to low reinforcement and low trial presentation together. The neurological data suggested that directly trained trials were the most sensitive to the experimental manipulations of reinforcement and trial presentation. Experiment 5.2 aimed to pinpoint the neurological process underlying the nodal number effect in equivalence class formation. RTs were a function of nodal number. Greater positivity of the P300 that is normally associated with categorization was found in 4-node than 1-node relations. No robust negativity 400 milliseconds after the target onset (N400) was found in either experiments (i.e. Experiment 5.1 and 5.2), contradicts Barnes-Holmes et al, (2005). In conclusion, the current thesis argued that nodal distance is a genuine effect that can not be explained as an experimental by-product.

Declaration and Statements

Declaration

This work has not been previously accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ....

Date ....29/06/2010........

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ....

Date ....29/06/2010........

Statement 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter library loan, and for the title and summary to be made available to outside organizations.

Signed ....

Date ....29/06/2010........

Table of Contents

Acknowledgement

It is always at this moment, we start to realise how lucky we are. I'd like to take this opportunity to thank my parents, without whose selfless support and belief in my ability, I would have spent the rest of my life as a teacher in a small school in my home town, never dreamt of the world outside, let alone attempted a Ph.D.

The completion of my Ph.D. marks the peak of my achievements so far, and will be cherished for the rest of my life, especially the six years spent with my supervisor, Louise; you are not merely a supervisor who helped me with her knowledge and guidance, rather, you are more a friend to me, the one who continuously encouraged, and inspired me at the times when completion of the thesis seemed impossible.

My special thanks to Rob in Ireland, whom I only met once, yet always offered patience, encouragement and support. I owe a debt to Louise, Natalie, and Carlo, for their patience and support on equipment handling, which made my first EEG experiment possible. I also thank Phil and Simon for their support and intriguing questions that keep me thinking. Furthermore, I can't image having completed this work without the support of the caring community in the department, especially Lisa, Neil, and the fellow Ph.D. students, Maria, Marcelle, Darren, Liv, Charlotte, Nic, Louise, Anita, J. J., Jordan, Claire, Alice, Thano, and Nigel, all of whom made every day of the journey special.

Equally, credit goes to my friend Barbara, and those back home, for their sympathetic ear and support that lift my spirit and spur me on. Even though I'm not religious, I do feel blessed that, when facing the greatest challenge of my life, I'm not alone. Thank you, Fi, for your love, for your tolerance, and for just being there with me.

Last, but not least, I thank all my participants, without their patience and effort I would not stand where I am now.

Research

Conference Publications/Presentations

2010

Dymond, S., **Wang, T.** (2010) *ERP correlates of relational learning II: testing a behavioural model of visual-visual priming.* Association for Behavior Analysis, San Antonio, TX, May 2010.

2009

Wang, T., McHugh, L., & Whelan, R. (2009) *The effect of nodal number on transitivity and equivalence relations.* Association for Behavior Analysis, Phoenix, AZ, May 2009.

2008

Wang, T., McHugh, L., & Whelan, R. (2008) *neurological correlates of semantic learning: A behavioural approach.* Welsh Institute of Cognitive Neuroscience (WICN), Swansea, September 2008.

Wang, T., McHugh, L., & Whelan, R. (2008) *A neurological measure of the effect of reinforcement on equivalence class formation.* European Association for Behavior Analysis, Madrid, Spain, September 2008.

2007

Wang, T. (2007). *An empirical account of delayed emergence of equivalence classes.* BPS Welsh Branch Student Conference, Swansea, Jun 2007.

Wang, T., Whelan, R., McHugh, L. (2007) *Empirical accounts of equivalence class formation.* Association for Behavior Analysis, San Diego, CA, May 2007.

2006

McHugh, L., Whelan, R., **Wang, T.** & Chatfield, C. (2006). *Examination of nodal-distance effects in equivalence class formation.* European Association for Behaviour Analysis, Milan, Italy, September 2006.

Whelan, R., McHugh, L., Thomas, C., & **Wang, T.** (2006) *Examination of nodal-distance effects in equivalence class formation.* Association for Behavior Analysis, Atlanta, GA, May 2006

Journal Articles

Wang, T., Charlotte, D., McHugh, L., & Whelan, R. (In press). Preserved nodal number effects under equal reinforcement. *Learning and Behavior*.

Wang, T., McHugh, L., & Whelan, R. (Invited to re-submit). A comparison of the nodal and discrimination accounts in equivalence formation. *Learning and Motivation.*

Academic Grants

McHugh, L., Whelan, R., & **Wang, T**. (2008).Semantic network growth: neural measures of the contribution of reinforcement. Welsh Institute of Cognitive Neuroscience (WICN), £2,200.

Tables and Figures

Tables

Figures

Illustrations

Chapter 1: Nodal Distance and Equivalence Class Formation

1.1 General Introduction

Language is pervasive in most if not all areas of life either publicly (i.e. talking) or privately (i.e. thinking). Humans are continuously describing, categorizing, relating, evaluating, writing about, reading about and thinking about, everything around them. The psychology of language originated from research in the area of linguistics (Bloomfield, 1914), an approach that focused primarily on structure (i.e. morphology and syntax), sounds (phonology), and meaning (i.e., semantics). Although these are important aspects of language this approach is narrow, ignoring the greater scope of language involved in higher cognitive functioning (e.g. decision making, problem solving, etc.).

Over the last five decades language has become of increasing interest to psychologists with the work of Skinner, 'Verbal Behavior' (1957) prompting much debate. According to Skinner (1957) an organism's public and private events (e.g. talking and thinking) are both forms of behaviour. From this perspective all behaviour is controlled by its antecedents and consequences based on a limited number of basic processes. Complex human affairs can be explained by these basic processes, without reference to additional information, such as the "mind", or "emotions", as these concepts are unnecessary metaphorical explanations. However, this behavioural position has been criticised largely due to the fact that it does not account for the generativity of many cognitive processes that are of primary interest to psychology (Chomsky, 1959). Nevertheless the search for a behavioural account of language acquisition continued and gradually became a central debate in modern behaviour analysis. A pivotal step in this regard was the emergent literature on "*stimulus equivalence*", a phenomenon that was believed to overcome the limitations Skinner had faced in providing a behavioural account of language acquisition (Sidman, 1994, Hayes, et al, 2001, also see Section 1.3.2 for more details).

The current chapter aims to provide readers with a basic understanding of behavioural principles and the commonly employed procedures used to demonstrate stimulus equivalence empirically. Historically, the behavioural literature on mediated associations provided pioneering research on stimulus relatedness, and thus the current chapter will follow that chronological order. Relevant research from cognitive psychologists aiming to provide evidence on semantic network formation will then be addressed. This will be followed by recent evidence from cognitive neuroscience and its overlap with current work in the area of stimulus equivalence and recent work in the area of behaviour analysis and semantic network growth.

1.2 Behavioural Principles and Their Applications in the Laboratory

The research methods, theoretical debates, and terminology used in the current thesis are inherited from animal studies in the behavioural tradition; therefore, it is important for the reader to have a basic background in behavioural principles and techniques before introducing the specific issues that will be addressed in the current thesis, therefore, basic behavioural principles are introduced in the next session.

1.2.1 Classical Conditioning

Classical or 'Pavlovian Conditioning' is a phenomenon that was originally demonstrated by Pavlov and his laboratory dogs. In these experiments the dogs salivated not only in the presence of food, but also in the presence of a neutral stimulus (i.e., a bell tone) that was consistently followed by the food. In this case, the food served as an Unconditional Stimulus (US) that always led to the secretion of saliva and this response served as an Unconditional Response (UR). When a bell tone was added prior to the presentation of food, the bell tone became a Conditional Stimulus (CS) that led to the Conditional Response (CR), the secretion of saliva, even in the absence of the food. The elicited CR in the presence of CS is referred to as "Classical Conditioning". The literature on classical conditioning has identified a number of important emergent patterns of behaviour surrounding this phenomenon. The sequence of the CS and US pairing can result in trace (the CS begins and ends before the US is presented), delayed (the US appears before the CS has disappeared), simultaneous (the CS and US coincide exactly) and backward (the CS follows the US) experimental alternatives to classical conditioning, even though not all of which work as effectively as classical conditioning. As classical conditioning is mostly if not all associated with simple reflexes, it has often been criticised for lacking the necessary depth to account for complex human behaviour such as language (Chance, 2008).

1.2.2 Operant Conditioning

Unlike classical conditioning, for operant conditioning to occur, behaviour must be emitted by an antecedent and be a function of its consequences, that is, the behaviour must be strengthened or weakened by its consequences. Operant conditioning is defined in terms of the three-term contingency, 1). the antecedent of the behaviour (e.g. red light), 2). the behaviour that occurs in that situation (e.g. lever press, and 3). the consequences of the behaviour (e.g. food delivery, Skinner, 1938). Catania (1998) identified three characteristics of reinforcement. First, the behaviour must have a consequence. Second, the behaviour must increase in its occurrence. Third, the increase in strength must be the result of the consequence. There are two types of reinforcement, positive and negative reinforcement. For positive reinforcement, the occurrence of a behaviour is increased by the presence of a stimulus called a "positive reinforcer". For example, a good score in an exam is a positive reinforcer, which must increase the probability of studying hard. In negative reinforcement, the occurrence of a behaviour is increased by the removal of a stimulus called a "negative reinforcer". For example, the removal of boring lecture slides could reduce boredom and thus increase attendance.

A number of phenomena have been demonstrated in the literature on reinforcement that warrant discussion. First, the rate at which behaviour occurs varies with the rate a behaviour is followed by a reinforcer. For example, the occurrence of a behaviour is increased if the reinforcer that followed is increased (Hammond, 1980). Second, the interval between a behaviour and its reinforcer affects the occurrence of a behaviour. For example, the shorter the interval is, the faster the behaviour emerges (Escobar & Bruner, 2007; Schlinger & Blakely, 1994; Lattal, 1995).Third, the magnitude and size of the reinforcer. For example, small reinforcers given frequently usually produce the occurrence of behaviour more fast than large reinforcers given infrequently (Schneider, 1973; Todorov, et al, 1984). When other variables are held equal, a large reinforcer is generally more effective than a small one (Christopher, 1988; Ludvig et al, 2007). Fourth, the occurrence of a behaviour is affected by the natural environment within which the behaviour occurs. For example, tasks that most resemble an animal's natural environment are more likely to induce certain behaviours (Hugdahl, 1995, 2001). Fifth, anything that makes a consequence more reinforcing necessarily changes the outcome of a reinforcement procedure. For example, the greater the level of deprivation of food, the more effective the reinforcer will be (Cotton, 1953; Reynolds & Pavlik, 1960). This phenomenon is explained as motivating operant behaviour (Michael, 1982, 1988, 2000; McGrill, 1999; Smith & Iwata, 1997). Sixth, previous learning experience with experimental stimuli can impact on speed of the occurrence of a behaviour (Sidman & Tailby, 1982). Seventh, the effects of reinforcing a behaviour will be very different if the behaviour also involves punishing consequences or if the reinforcers are simultaneously available for other kinds of behaviours (Herrnstein, 1970). Reinforcement is the only important variable in the production of verbal behaviour, according to Skinner, verbal behaviour is shaped and maintained by social interaction with other people in the environment (Skinner, 1957).

1.2.3 Discrimination

Discrimination is the tendency for behaviour to occur in situations that closely resemble the situation in which the behaviour was learnt but not in situations that differ from it. Any procedure used for establishing a discrimination is called discrimination training (Chance, 2008). Discrimination procedures are widely used in both the classical and operant conditioning literatures. In classical discrimination training, one reinforced CS (CS+) (e.g. tone) is paired with a US (e.g. food delivery), and a non reinforced CS (CS-) appears alone (e.g. red light), the desired behaviour (e.g. salivation) will occur in the presence of the CS+, but not in the presence of the CS- (Pavlov, 1927). In operant discrimination training, a stimulus (S+) is associated with a reinforcing consequence, whilst a stimulus (S-) is associated with no reinforcing consequence. This kind of discrimination contingent upon a behaviour's consequence is also known as a conditional discrimination. There are three typical S+ and S- presentation formats used to produce conditional discrimination, these are, successive, simultaneous and Match to Sample (MTS) procedures. In successive discrimination training, the S+ and S- are presented alternately. In simultaneous discrimination training, the S+ and S- are presented at the same time. In MTS training, the task is to select the reinforced comparison stimulus (C+) from two or more comparison stimuli (C+ and non reinforced comparison stimulus (C-s)) to match a standard stimulus (the sample S+). There are two types of matching, identity and arbitrary matching. In identity matching, C+ resembles a certain aspect of S+ (e.g. shape, colour, font, size), whilst in arbitrary matching the association between C+ and S+ was established by reinforcement. The latter will be used in all empirical chapters

(Chapter 2 – 5) presented in the current thesis. A variation of MTS procedure is asked to select the C- from comparisons that differ from the sample, this is called a mismatching procedure or oddity matching.

However, these procedures often result in large number of errors. One way to reduce errors would be to vary the delivery of reinforcers with the target behaviour, this method in training a discriminative performance is known as the "Differential Outcomes Effect" (DOE) (Peterson & Trapold, 1980; Trapold, 1970; Miyashita, et al, 2000). For example, discriminative behaviour (e.g. press red key, press green key) reaches maximum performance levels if different behaviours are reinforced in different ways (e.g. red key was discriminated by chocolate, green key was discriminated by candy). This method will be further discussed in the context of stimulus equivalence (see Section 1.3.3.1).

1.2.4 Generalisation

A related but different behavioural pattern is referred to as generalisation. Generalisation is the tendency for a behaviour to occur in situations different from the one in which the behaviour was learned (Chance, 2008). For example, in the infamous Watson and Rayner (1920) study, little Albert leaned to fear the rat. This fear generalised to a rabbit, raw cotton, and a Santa Claus mask after the experiment. That is, Albert showed a generalised fear from the white rat to other white furry objects. However, generalisation does not always emerge (Ducharme & Holborn 1997). In the behavioural learning literature, several studies have demonstrated that the degree of generalisation of a learned behaviour in the presence of stimuli other than the trained stimulus is based on the level of similarity between stimuli and the trained stimulus (Hovland, 1937; Guttman & Kalish, 1956; Guttman, 1963). Hence, learned behaviour is most likely to appear in situations that closely resemble the training situation.

The trained stimulus and generalized stimuli must share similar characteristics to allow generalisation to occur, those characteristics can be both physical and abstract. A stimulus's physical characteristics can include colour, size, shape, pitch, loudness, to name but a few. Whereas a stimulus's abstract characteristics are mainly studied in the generalization of semantic information (Razran, 1939; Lacey et al, 1955). That is, learned behaviour is more likely to generalise amongst words that are semantically related. Given the importance of the generalisation of semantics for the focus of the current thesis a summary of the semantic generalisation literature will follow.

1.3 Semantic Generalisation and Stimulus Equivalence

Behavioural research on semantic generalisation paved the way for what was later referred to as stimulus equivalence, therefore semantic generalisation will be discussed first.

1.3.1. Mediated-associations

Semantic generalisation, sometimes refers to mediated transfer, mediated association, or mediated generalization (e.g. if the association between word A and B is the result of associations of both A and B to a third word C, this is referred to as a mediated association). It was first described by Cofer and Foley (1942) as depending not on physical similarities amongst stimuli but on arbitrary stimulus equivalence that has been established by previously conditioned behaviour. Hence, semantic generalisation is a result of conditioning. However, the verification of this assumption is not that straightforward. Early research in this area has been limited to what is referred to as the "three stage transfer", a term that eventually lost popularity after the failure to demonstrate any four stage mediated generalization (Jenkins & Palermo, 1964; Jenkins, 1963; 1965) through systematically manipulating the paired-association paradigm typically used in three stage mediated generalization (i.e., Underwood, 1949). In a paired-association paradigm, pairs of words are presented successively. The first word is the stimulus, the second word is the response and the participant has to associate each response with its stimulus. The number of correct responses and response time are compared with a control group who had not undergone mediated training. Hence, participants who demonstrate mediated associations learned faster and were more accurate in the presence of a new list of paired associated words. For instance, participants exposed to a SUN – EARTH, pair would learn to say EARTH when SUN is presented. In the paired-association literature, the association between stimulus and response was illustrated using three stages: if A elicits B (the first stage); and A elicits C (the second stage), then, B will tend to elicit C and C will tend to elicit B (the third stage).

Jenkins (1963) identified four types of mediated association based on observation: simple chains (A →B→C), reverse chains (A←B←C), stimulus equivalence (A→B←C), and response equivalence (A←B→C). As we shall see, the paired-association literature ended at this point. Jenkins (1963) conducted a systematic analysis of the functional equivalence amongst stimuli in mediated generalization through 16 paired-associate paradigms in order to examine the emergence of a four stage transfer, which was considered critical in verifying mediational accounts (Cofer & Foley, 1942) For instance, Learn A →B (the first stage), learn C→B (the second stage), then learn A→D (the third stage), test C→D (the forth stage). It was assumed that A and C become functionally equivalent after the first two stages, when a new response is learned to A in the third stage, it should tend to occur in the presence of C in the test stage. In another paradigm, if a participant learns A→B, C→B and then D→A and is tested on D→C, functionally equivalent stimuli must become functionally equivalent responses to produce the DC association. Unfortunately, none of these 16 paradigms successfully yielded the four stage mediated transfer. Sidman (1994) highlighted a methodological problem across all paired-association paradigms. He argued that the successful learning of the CD relation in the first example required the maintenance of AB, CB performance (also known as baseline relations) during the test. However, in a typical third stage of paired-association training, AB relations were likely to undergo extinction due to a lack of reinforcement, whist in the test stage, CB relation were likely to go into extinction. Therefore, a partial solution to maintaining the baseline relations could be achieved by re-exposing participants to the lists that they had learned in the first two stages (Grover, Horton, & Cunningham, 1967; James & Hakes, 1965), but the intactness of the baseline relations in the test stage was still ambiguous, which resulted in a misinterpretation of the lack of mediated associations.

1.3.2. Stimulus Equivalence

1.3.2.1 Human Evidence

The first mediated association that was demonstrated in a conditional-discrimination paradigm was discovered incidentally by Sidman (1971) while teaching reading comprehension to a retarded boy, who was unable to read printed words orally or with comprehension, but could match spoken words to pictures and could name pictures. Sidman taught him to match spoken words to printed words using a conditional-discrimination paradigm as described earlier (see section 1.2.1.3). Surprisingly, the boy demonstrated the ability to read (matching the printed words to pictures) and oral reading (naming the printed words). To illustrate: known A (auditory words) →B (picture), teach A→C (visual words), test B→C, and C→B; known D (naming) →B, test D→C. Therefore, B and C were demonstrated to be functionally equivalent to each other. However, in order to make sure the conditional-discrimination procedure did not involve the same difficulties as the paired-association procedure, the demonstration of four stage mediated transfer was necessary.

Sidman and Tailby (1982) trained eight healthy 5- to 7-year-olds on three three-member stimulus classes first (A1B1C1, A2B2C2, A3B3C3) using a conditional-discrimination procedure, that is, teaching A1B1, A1C1, A2B2, A2C2, A3B3, and A3C3, then test emerged relations that have never been taught before (B1C1, B2C2, B3C3, C1B1, C2B2, and C3B3). After successful demonstration of those emerged new relations, a forth stimulus was added to each functional equivalence class (stimuli D), that is, teaching D1C1, D2C2, and D3C3, then test D1B1, D2B2, D3B3, B1D1, B2D2, and B3D3 emergent relations. Six children demonstrated these new emergent relations, and also demonstrated proficiency across three four-member stimulus classes by demonstrating AD and CD relations whilst keeping the original BC and CB relations intact.

There were some distinctive features of this study which warrant discussion here. First, the stimuli were both dictated Greek letter names (stimuli A) and printed Greek letters (stimuli B, C and D) to ensure the emergent responses were only produced via discriminative training rather than pre-experimental learning. Second, unlike paired-association procedures, baseline relations were closely monitored during testing for emergent relations (i.e., a reinforcement fading procedure), and the comparison stimuli in training and testing were restricted to only one type (e.g., B) of stimuli across four classes (e.g. B1, B2, B3 and B4). Therefore, no extinction of baseline relations described in the paired-association procedure would occur. Third, participants were taught to name stimuli A, then test naming in the presence of stimuli B, C, and D.

The successful generation of equivalence from three-member to four-member classes is very important both empirically and conceptually. It marked the start of a wealth of research into the equivalence phenomenon, here; the term equivalence instead of mediation or mediated transfer was defined purely in terms of mathematics in order to describe the emergence of reflexivity, symmetry, and transitivity relations that had not been directly taught after discrimination training.

Equivalence classes can be established by training a minimal number of relations between the stimuli in a group or class. For example, if the group of stimuli consisted of the letters A, B, and C, an equivalence class could be established by training two, two-term relations between AB, and BC, using a conditional discrimination paradigm, or respondent condition procedures (Sidman, 1971, 1994). If a class has been established, many new emergent relations are formed between the stimuli that had not been taught directly. There are four types of emergent relations (see Bush, Sidman, & de Rose, 1989), and examples of each are given for the group (ABC) described above: (i) reflexive relations (A → A, B → B, C → C), (ii) symmetrical relations (B → A, and C → B), (iii) transitive relations (A → C), and (iv) equivalence relations (C → A). If all of the emergent relations control responding, then the group of stimuli can be said to function as an equivalence class (Sidman, Kirk, & Wilson-Morris, 1985), and the stimuli are fully substitutable for one another (Sidman, 1990, 1994).

The equivalence phenomenon has been replicated and extended with human participants of varying ages, some as young as 2-years old (Beasty, 1987; Lowe & Beasty, 1987; Dugdale & Lowe, 1990; Sidman, Kirk, & Willson-Morris, 1985; Smeets, Roche, & Barnes-Holmes, 1997; Smeets, Barnes-Holmes, & Cullinan, 2000; Goyos, 2000). Equivalence also demonstrated in healthy participants and participants with vary severity in retardation (Sidman, 1971; Sidman & Tailby, 1982; Sidman, Cresson, & Willson-Morris, 1974; Sidman, Willson-Morris, & Kirk, 1986; Devany, Hayes, & Nelson, 1986; Carr & Felce, 2000; Carr, et al., 2000). Different stimulus modalities also produced equivalence performance, for example, visual stimuli (pictures, syllables, nonsense words, letters, objects, colour and Mandrin characters) (Hayes, Tilley, & Hayes, 1988; Leslie, et al., 1993; Stromer & Osborne, 1982; Lazar et al., 1984; Sidman, Kirk, & Willson-Morris, 1985), auditory stimuli (spoken words, music, none-sense syllables) (Sidman, 1971, Sidman & Tailby, 1985; Dugdale & Lowe, 1990; Beasty, 1987; Stromer, Mackay, & Remington, 1996; Hayes, Thompson & Hayes,1989; Dube, Green, and Serna, 1993). Although visual and auditory stimuli are the most commonly used, there were other studies that have demonstrated equivalence with drug related stimuli (De Grandpre, Bickel, and Higgins, 1992), gustatory stimuli (Hayes et al, 1988), tactile stimuli (Bush, 1993), olfactory stimuli (Annett & Leslie, 1995) and equivalence is not restricted to one stimulus modality (Sidman 1971; Dugale & Lowe, 1990; Lowe & Beasty, 1987).

Importantly, after an equivalence class is established, and a function is established for one member of the class, that function may transfer to other members of that class in the absence of explicit training (Fields, et. al., 1993, 1995, Fields & Moss, 2007; Fields & Watanabe-Rose, 2008). For example, if A, B and C are members of an equivalence class, and A acquires anxiety eliciting functions through pairing with shock, then B and C may acquire a similar function without being similarly associated with shock (Augustson & Dougher, 1997). This phenomenon is referred to as *transfer of function.*

Behavioural researchers have demonstrated the derived transfer of a variety of stimulus functions, including self-discrimination (Dymond & Barnes, 1995), aversive respondent-eliciting functions (Dougher, Auguston, Markham, Greenway, & Wulfert, 1994), sexual arousal functions (Roche & Barnes, 1997; Roche, Barnes-Holmes, Smeets, Barnes-Holmes, & McGeady, 2000), avoidant evoking functions (Augustson & Dougher, 1997; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007), self-reported arousal functions (Smyth, Barnes-Holmes & Forsyth, 2006), mood-generating functions (Barnes-Holmes, Barnes-Holmes, Smeets & Luciano, 2004; Cahill, Barnes-Holmes, Barnes-Holmes, Rodriguez-Valverde, Luciano, & Smeets, 2007), preference functions (Barnes-Holmes, Keane, Barnes-Holmes, & Smeets, 2000; Smeets & Barnes-Holmes, 2003), and self-efficacy functions (Gutiérrez-Martínez, Luciano-Soriano & Valdivia-Salas, 2005), through stimulus classes in a range of experimental contexts (Barnes & Keenan, 1993; Bush, Sidman, & de Rose, 1989; Hayes, Brownstein, Devany, Kohlenberg, & Shelby, 1987; de Rose, McIlvane, Dube, Galpin, & Stoddard, 1988; Wulfert & Hayes, 1988, Dymond & Rehfeldt, 2000). Transfer of function or Transformation of functions will be revisited in Section 1.3.3.3.

1.3.2.2. Animal Studies

One example of stimulus equivalence in the animal conditioning literature is a phenomenon referred to as "transitive inference" (TI). Specifically, given a relation between A and B, between B and C, an animal would be expected to *infer* the AC relation without explicitly exposed to the stimulus pair. There is a growing body of research demonstrating transitive inference (TI) in a number of species (see

Vasconcelos, 2008 for a review). The typical method for assessing TI in animals involves standard operant conditioning, like those used in the paired-association literature, with the major difference being that responses emitted by each stimulus are presented with both reinforcing and non-reinforcing consequences. For example, in a 5-term linear training paradigm, stimulus pairs AB, BC, CD, and DE were introduced in a linear fashion, that is, the animal learns to discriminate between stimulus A and B first by responding (e.g. by pecking a lever) to stimulus A (resulting in reinforcement: A+), responding to stimulus B (not resulting in reinforcement: B-); then they learn to discriminate between stimulus B and C by responding to stimulus B (resulting in reinforcement: B+), responding to stimulus C (not resulting in reinforcement: C-); then the discrimination between the CD stimulus pair is learned through the same differential reinforcement described in other stimulus pairs (C+, D-). Finally, discrimination between DE are learned through the same process (D+, E-). This type of training is illustrated as follow: A+B-, B+C-, C+D-, D+E-, "+" marks reinforcement, whist "-" marks non reinforcement. In the test session, stimulus pair BD are presented. Transitive inference is demonstrated if a preference over stimulus B rather than stimulus D emerges by responding toward B more often than to D. The underlying rational is as follows, as stimulus B, C, and D receive an equal number of reinforced and non reinforced trials, responding to B over D should not come from differential reinforcement, but be analogues to an inference, such as more than, A>B>C>D>E.

This serial training paradigm, however, is different from the conditional-discrimination procedure, which generates a four-term contingency (Sidman, 1994, we will revisit this in the next section). Thus, responses in a TI training paradigm are not conditional upon the relationship between the conditional stimulus and the discriminative stimulus, but merely a simple response discrimination towards reinforcement, therefore, the three-term contingency is also called a simple discrimination. However, how could a response controlled simply by its reinforcement history account for responses controlled by the discrimination between stimulus associations (equivalence performance, see Section 2.3.1)? In another words, how do we know a pigeon's pecking behaviour in the presence of a red light is not because it has been reinforced to do so, but it "knew" the association between red light and say blue light? Of course, one could argue that by presenting B and D in the test of TI, eliminates the effect of reinforcement, as B and D received the same amount of reinforcement/non reinforcement. Another issue in the TI literature is that it only involves the test for a preference between the B and D stimuli which is not in line with the definition of equivalence. Could pecking in the presence of B over D (TI performance) be due to other undefined variables? One potential explanation for this preference might be the way the stimuli were presented. Specifically, the fact that the order sequence in which each stimulus is presented can affect preference towards these stimuli is well documented. The serial position effect, overarching effect, end-arching effect, etc., have all been reported to produce stimulus preferences (Wynne, 1997; Semann, et al., 1996). Some even demonstrated the serial position effect independent of the training structure in which all discriminations were trained concurrently (Wynne, 1997; Semann, Delius, & Wright, 1996; For a detailed account of the debate see Vasconcelos, 2008). Nevertheless, the methodological discrepancies (mainly, the test procedure) and conceptual issues between the TI and equivalence literature render comparisons between these two theories difficult.

A paradigm from the animal literature that closely resembles stimulus equivalence preparations was developed from a modified associative learning paradigm, in which specific reinforcement is associated with a particular consequence. This differential reinforcement procedure (described in Section 1.2.3. as differential outcome effect) enables a test for the three properties of equivalence (reflexivity, symmetry, and transitivity) (see Sidman, 1994, p382-383;) in the three term contingency, this is known as "acquired equivalence" (AE) in the animal literature (c.f. Schusterman & Kastak, 1993; Manabe, Kawashima, & Staddon, 1995; Ward-Robinson & Hall, 1999; see also Urcuioli, 2001 for a review). However, as Urcuioli (2001) and others (c.f. Saunders, Williams, & Spradlin, 1996; Zentall, 1998) have pointed out "there are equivalence phenomena readily exhibited by humans that have little, if any, known counterpart in pigeons and other animals." (p. 16) This assumption is based on the fact that emergent relations mostly observed in the AE literature has two distinctive experimental prerequisites, that is, the differential reinforcement procedure and successive discrimination during baseline training (see Section 1.3.3.4. for a discussion on experimental procedures). Therefore, it is not clear whether AE is merely a result of these experimental procedures, or equivalence performance as their human counterpart. Although according to Sidman's definition of equivalence, there is no reason why an animal cannot successfully form equivalence relations, an equivocal demonstration with precise experimental control is needed before any conclusion can be made. After all Sidman's analysis of equivalence remains controversial (Hayes, 1989; Pilgrim & Galizio, 1995; Saunders & Green, 1992; Horne & Lowe, 1996). The section that follows aims to provide the reader with a background to the stimulus equivalence debate.

1.3.3. The Origin of Equivalence

1.3.3.1 Sidman's Contingency Theory

Sidman's theory of equivalence evolved through continuous observation. In his 1994 book, equivalence is presented as a direct outcome of reinforcement contingencies based on Skinner's three-term contingency theory (1953). As described in Section 1.2.3, the discriminative stimulus, response emitted by that discriminative stimulus, and response consequence, forms a three-term analysis unit in operant conditioning. Sidman argued that the number of terms included in the analysis acts as an additional dependent variable and is a result of the type of training procedure employed. For example, in a conditional-discrimination procedure, conditional stimuli are added into the three-term unit, and ultimately extend the analysis unit into a four-term unit. That is, a response is conditional upon the association between discriminative stimulus and conditional stimulus. According to Sidman (1994), conditional-discrimination procedures produce two distinctive outcomes: one is the 4-term contingency, the other is equivalence performance. Therefore, equivalence according to Sidman (1994) is the primary unit of language acquisition, and needs no further explanation. Sidman (1994), admitted that the mathematical definition of equivalence is mainly descriptive rather than explanatory. The mathematical definition of equivalence and its unexplainable nature have been the subject of various criticisms (Hayes, 1989; Horne & Lowe, 1996; Saunders & Green, 1992).

In response to these criticisms, Sidman summarised and extended his 1994 work, and proposed an empirically verifiable account of the origins of equivalence (2000). According to this proposal, contingencies of reinforcement produce two outcomes: the *n*-term contingency, that equivalence performance is not restricted to the four-term contingency unit, for example, a five-term contingency might account for second-order conditional discriminations or the rule-governed behaviour that is commonly observed outside the laboratory; the other outcome of contingencies of reinforcement is equivalence relations which can be described as "ordered pairs of all positive elements that participate in the contingency" (Sidman, 2000, p128). This account emerged from observations that when one common reinforcer and response participate in the conditional-discrimination procedure and join with other stimuli within the contingency this forms an extended equivalence class (Dube, et al., 1987; 1989; McIlvane, et al., 1992). He argued that, despite the *n*-term unit, contingencies of reinforcement also establish two conflicting outcomes, that is, the extended equivalence class that combines stimuli from both classes is generated by one common reinforcer and responses to this extended equivalence class should emerge before being broken down to smaller class discriminations (Sidman, 2000). Sidman outlined various experimental paradigms based on whether the class discrimination is reinforcement or response specified. If tests for these performances yield positive results, this would support the contingency theory.

The argument that differential outcomes facilitate mediated associations is well documented in the literature (see section 1.2.3). For example, Minster and colleagues (2006) provided a direct test of one of the experimental paradigms proposed by Sidman (2000). In Experiment 1, four three-member equivalence classes were established via different stimulus-reinforcement relations. Two of them were established with class-specific reinforcers, R1 and R2 (A1B1C1R1, A2B2C2R2), the other two classes were established with one common reinforcer, R3 (A3B3C3R3, A4B4C4R4). The study aimed to determine whether R1 and R2 are functionally equivalent to corresponding equivalence class members (e.g. A1, A2); whether R3 drops out of the extended equivalence class to allow class-specified responses to emerge (e.g. A3B3C4, A4B4C4) through a many-to-one MTS paradigm. Findings from this work indicated that all three reinforcers participated in corresponding equivalence classes (A1B1C1R1, A2B2C2R2, A3B3C3R3, A4B4C4R3) regardless of different stimulus-reinforcement relations. Experiment 2 aimed to control for the exclusion of interclass relations in Experiment 1 with the addition of an interclass relations test at the end, that is, samples from class 3 were presented with comparisons excluding members from class 3 and samples from class 4 were presented with comparisons excluding members from class 4. It was predicted that successful performance on interclass relations would be a result of the common reinforcer R3 that has been associated with class 3 and 4. This is exactly what they found. The authors suggested that common reinforcers can participant in class-specified equivalence, thus refuting Sidman's (2000) account that a common reinforcer has to drop out to allow class-specified equivalence to emerge.

Sidman (1994) argued that stimuli may be members of multiple classes because class membership may be controlled by contextual cues (Bush, et al, 1989; Gatch & Osborne, 1989; Meehan & Fields, 1995), thus, the removal of class-consistent comparisons in Minster et al.'s (2006) Experiment 2 might establish some implicit contextual cues that facilitate the stimulus-reinforcer relations. Minster and colleagues not only acknowledged the possibility of contextual control in accounting for their findings, but also suggested that repeated training and testing of equivalence results in differential reinforcement in stimulus-stimulus and stimulus-reinforcer relations might also account for establishing stimulus-stimulus relations prior to stimulus-reinforcer relations.

1.3.3.2 Naming

One major criticism of Sidman's equivalence is that equivalence performance is not the primary unit of language acquisition, rather naming provides the basis for equivalence performance (Dugdale & Lowe, 1990; Lowe, 1986; Horne & Lowe, 1996). These researchers have demonstrated that naming can facilitate equivalence class formation (Green, 1990; Lipkens, Hayes, & Hayes, 1993; Sidman, et al., 1986). Horne and Lowe (1996) adopted an intensive developmental account of how naming is acquired at an early age, maintained, and generalized in a complex interaction both verbally and non-verbally (e.g. visual, tactile features) between the caregiver and the child. They argued that, unlike the distinctive role between speaker and listener postulated by Skinner (1957), the child learned to name through not only a passive listener's position, but also an active speaker to him/herself during interactions with her environment. According to them, naming can be defined as a higher order bidirectional behavioural relation and includes three features. First, it combines conventional speaker and listener behaviour within the individual, as described earlier. Second, it does not require reinforcement to allow the emergence of new behaviour, as multiple exemplar training establishes contextual control over behaviour. Third, it applies to classes of objects and events. Therefore, explaining the massive expansion of vocabulary in early age.

However, Sidman (1982) argued that researchers who assume naming is the basic unit of language acquisition are influenced by the same assumption of stimulus-response association that is proposed in the mediated associations literature (Cofer &

Foley, 1942). That is, in testing for emergent relations, responses learned to the original stimulus will occur first, and then different responses will occur in the presence of other members of the equivalence class. He (1982) further suggested that the response-response equivalence amongst stimuli is not a prerequisite for equivalence performance; rather, a direct stimulus-stimulus relation would be sufficient to establish this performance. Empirical studies demonstrated that accurate auditory-visual matching and auditory-visual equivalence can be demonstrated in the absence of consistent naming (Sidman & Cresson, 1973; Sidman, Willson-Morris, & Kirk, 1986; Saunders & Green, 1996). Nevertheless, Sidman later (1994) dropped this distinction between stimulus and response, arguing that all stimulus pairs, no matter whether it is a stimulus-stimulus pair, stimulus-response pair, response-response pair, stimulus-reinforcer pair, or a response-reinforcer pair, would be able to establish equivalence in light of evidence that response and reinforcer can also participate like stimuli in an equivalence class (Dube, et al., 1987; 1989; McIlvane, et al., 1992).

In response to these critiques, Horne and Lowe (1996) argued that the lack of consistent naming in equivalence studies was due to the procedure used, that is, all the evidence was found in post-experimental naming tests after the participant completed the conditional-discrimination procedure, the results obtained could be very different from those found through the conditional-discrimination procedure (Dugdale & Lowe, 1990), and suggested that naming can be established even without consistent naming. They also pointed out that there is almost no way to generate absolute abstract stimuli free from intra-stimuli relations. In return, they criticised the removal of the stimulus and response distinction, viewing it as the removal of the distinction between the behaviour and its environment. However, Sidman (2000) insisted that the responses participated in the reinforcement contingency functionally identical to other stimuli, he argued that the defined discriminative response is naming. Finally, the lack of general evidence of equivalence in animal studies (Sidman, et al., 1982; Hogan & Zentall, 1977; Homes, 1979, Rodewald, 1974; Dymond, Gomez-Martin, & Barnes, 1996) supports naming, a distinctive feature in humans, as the primary unit of verbal behaviour (Horne & Lowe, 1996). Not very long after the conception of naming as a theory of equivalence formation, evidence from the AE literature lead Urcuioli (2001) to conclude that the ability to name is not a prerequisite for AE performance, instead, it is mostly a reflection, rather than foundation process. However, as stated in Section

1.3.2.2., consideration of methodology and conceptual applications of equivalence to animal species needs to be fully explored before any conclusions can be drawn.

1.3.3.3 Relational Frame Theory

Extending on from Sidman's theory of stimulus equivalence, Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001) suggests that success on equivalence tests is a result of prolonged exposure to the contingencies of reinforcement operating in the verbal community (i.e., the multiple exemplar training). RFT accounts for equivalence as a type of *arbitrarily applicable relational responding* (otherwise referred to as *relational framing*) referred to as a relation of "Coordination" (Barnes & Holmes, 1991).

Human and nonhuman species are easily taught to respond to non arbitrary relations between stimuli, such as "smaller than", "taller than", "darker than" and "coordinate to". However, verbally able humans can learn to respond to relations between stimuli where these relations are not defined by the physical form of the stimuli, but by contextual cues (e.g. point to, orientate to) due to a prior history of learning. As only contextual cues (regardless of the physical features of the stimuli) are required, such relations are arbitrarily applicable to any event. According to RFT, arbitrarily applicable responding (relational framing) shares three common defining features, referred to as mutual entailment, combinatorial entailment and the transformation of stimulus functions, responses emergent based on these features were known as derived stimulus relations.

Mutual entailment is similar to the concept of symmetry in the equivalence literature and describes the relations that pertain between two stimuli or events (Sidman, 1992). That is, if explicit relations are established between two stimuli A and B, then a relation between B and A may also be derived (e.g. if A=B then B=A). From a relational frame perspective, it is important to emphasise that not all mutually entailed relations are equivalent (e.g. if A>B then B<A), as would be the case with the concept of symmetry. The particular relation between events is controlled by contextual stimuli.

Combinatorial entailment describes the relations that pertain among three or more stimuli. For example, from explicitly trained relations such as A>B and B>C, one can derive relations between A and C and between C and A, such as A>C and

C<A. Within the equivalence literature, combinatorial entailment is termed *transitivity* (A=C) and *equivalence* (C=A) (See Hayes, 1994, p.11).

The term *transfer of stimulus function* is used to describe changes in stimulus functions that result from their participation in relational frames. For example, in a specific context, if stimulus A is related to stimulus B (e.g. by opposition), and A is given a psychological stimulus function, the functions of B may be transferred in accord with the specified relation to A. Consider the following example taken from Hayes, Fox, Gifford, Wilson, Barnes-Holmes, and Healy (2001). Imagine a person who is trained to select a B stimulus as the "opposite" of an A stimulus. If the A stimulus then has a *punishing* function attached to it, RFT would predict that the B stimulus would then have derived *reinforcing* functions because of its "opposite" relation to the punishing A stimulus (see Dymond & Barnes, 1995). Evidence of function transfer was also demonstrated in "more than", and "less than" relational frames (Whelan, Barnes-Holmes, and Dymond, 2006).

However, much controversy has surrounded RFT since its conception due to its similarity to the concept of equivalence and verbal behaviour, indeed many behaviour analysts have criticised RFT as old medicine in a new bottle (Burgos, 2003; Palmer, 2004ab). Hayes and Barnes-Holmes (2004) argued that RFT not only provides an explanatory account of the origin of equivalence, but more importantly, it offers a comprehensive behaviour analysis of language acquisition that is verifiable by empirical tests, therefore distinguishable from Skinner's verbal behaviour.

1.3.3.4. Delayed Emergence and Nodal Distance

Another issue in the equivalence literature is that the predicted performance often requires repeated testing, as described in Minster et al. (2006). This delayed emergence of a new conditional discrimination poses a challenge to Sidman's contingency account of equivalence performance (1994, 2000). Sidman was well aware of this issue and admitted that conditional-discrimination training is not sufficient to establish equivalence performance, rather the MTS test itself gives rise to the emergent relations. This might partially account for the lack of 4-member equivalence in the mediated-association paradigm, and the lack of equivalence in the animal studies discussed earlier (James & Hake, 1965). In facing this difficulty, Sidman argued that contextual control emerged during the MTS test and facilitated the equivalence performance. However, contextual control as a description of equivalence class formation can raise problems for Sidman's account of equivalence. Adding the additional concept of "C*ontextual Control*" may trivialise any attempt to search for the real processes underlying equivalence formation. For instance, participants' verbal ability is argued to be the primary unit in verbal behaviour (see Section 1.3.3.2 for more detail).

Sidman's account of equivalence also emphasised the interchangeability of class membership, whereas it is often the case that delayed emergence of new conditional discriminations is not simultaneous, but follows a systematically ordered pattern, known as "*Nodal Distance*". A *node* has been defined as a stimulus that is linked by training to at least two other stimuli (Fields & Verhave, 1987; Fields, Verhave, & Fath, 1984) and the number of nodes that link any two stimuli in a set of trained conditional relations is described as the *nodal number* (Sidman 1994). For example, a 5-member class (A, B, C, D, and E) contains six 1-node relations (e.g., B-D, with C as the node), four 2-node relations (e.g., B-E, with C and D as nodes), and two 3-node relations (e.g., A-E, with B, C, and D as nodes). Several studies have reported that the probability of successful emergence is a function of nodal number, that is, response time is increased, and response accuracy decreases as nodal number increases (Bentall, Jones, & Dickens, 1998; Dube, Green and Serna, 1993; Fields, Adams, Verhave, & Newman, 1990; Kennedy, 1991; Lazar, 1977; Rehfeldt & Dymond 2005; Spencer & Chase, 1996; Wulfert & Hayes, 1988). Recently, there were intensive debates arguing whether nodal distance is genuine or just an experimental by-product (Fields et al, 1990; Sidman, 1994; Kennedy, 1991; Saunders & Green, 1999; Imam, 2001, 2006). This has lead researchers to focus on systematically manipulating training structure and protocols (Saunders & Green, 1999; Adams, Fields, & Verhave, 1993; Fields, et al, 1997) in order to examine their impact on equivalence formation and the relatedness of class members.

The term *training protocol* refers to the sequence of conditional discriminations presented in baseline training and testing (Field, et al, 1993; Imam, 2006). Three training protocols that are commonly used in equivalence studies are simple-to-complex (STC), complex-to-simple (CTS), and simultaneous protocol (SP). In STC, one baseline relation (AB) is trained followed by a test for symmetry (BA), then the new baseline relation (BC) is introduced and symmetry (CB), transitivity (AC) and equivalence relations (CA) are tested sequentially. In contrast, the test of equivalence relations is prior to the test of symmetry and transitivity relations after baseline training in CTS. As its name suggests, all emerged relations are tested simultaneously in one mixed block after baseline conditional discriminations are simultaneously trained in SP. The term *training structure* refers to the arrangement of linking stimuli presented in baseline training (Saunders & Green, 1999). For example, a linear training structure or serial training structure involves training A-B, B-C, and C-D, whereas a one-to-many structure or sample-as-node involves training A-B, A-C, and A-D, many-to-one structure or comparison-as-node, involves training B-A, C-A, and D-A. There was some evidence that SP alone has higher difficulty in producing emergent stimulus relations compared to STC and CTS (Fields et al., 1995, 1997). STC is more likely to produce emergent stimulus relations than CTS (Adams et al, 1993). Comparison-as-node training seems to yield better positive results on equivalence performance than sample-as-node training structure in children, adolescents and retarded adults (Spradlin & Saunders, 1986; Saunders et al, 1999). In contrast, one study with healthy adults found that sample-as-node training was more likely to produce three-member equivalence classes than comparison-as-node training (Arntzen & Holth, 1997). In addition, serial training results in differential reinforcement and has been argued to be responsible for nodal number effects (Fields, et al., 1997; Imam, 2001, 2006). Other factors, such as control of negative reinforcement (Sidman, 1994) and history of conditional discrimination on MTS training (Saunders & Green, 1999) have also been cited as responsible for the emergence of nodal number effects in equivalence classes. The first two empirical chapters in the current thesis aimed to test two of these assumptions by systematically manipulating equivalence training structures.

Other researchers (Fields, et al., 1993; 1995; Fields & Moss, 2007; Fields & Watanabe-Rose, 2008; Kennedy, 1991; de Rose et al, 1988; Bentall, et al., 1998) argue that the nodal number effect is genuine and disregard the experimental manipulations as responsible for the emergence of this phenomenon, suggesting that the strength of relatedness amongst equivalence class members varying across experimental procedures, is not sufficient to conclude the nodal number effect is a experimental by-product. For example, many studies (Fields, et al., 1993; Sidman 1994; Dickins, et al., 1993; Kennedy, 1991; Bentall et al., 1993) reported that nodal number effects were observed in the MTS test of emergent relations, but disappeared with repeated testing. Accuracy of responding to unreinforced probes for emergent relations has been the most common measure of the relatedness among stimuli.

Additionally, supplemental measures of emergent relational responding may shed more light on the nature of the relations among stimuli (Dymond & Rehfeldt, 2001). Several researchers have reported that response speed (RS) was a function of nodal number (Bentall, Dickins, & Fox, 1993; Bentall et al., 1998; Holth & Arntzen, 2000; Spencer & Chase, 1996; Wulfert & Hayes, 1988) even when accuracy remained intact. Fields, Landon-Jimenez, Buffington and Adams (1995; see also Fields, Adams, Verhave, & Newman, 1993) adopted another alternative measure of stimulus relatedness, namely a *transfer of function* test. In this study, two 5-member equivalence classes were trained using a protocol that ensured equal reinforcement across trial types. All (i.e., 12) participants passed baseline discriminations. However, only two participants formed equivalence classes. After these two participants demonstrated the formation of equivalence classes, new responses were trained to the end (i.e., A and E) stimuli in each group. Transfer of function was measured in terms of the relative frequency with which responses trained to A and E stimuli were evoked by all stimuli in both classes. In general, transfer of function was an inverse function of nodal number.

A recent study by Fields and Watanabe-Rose (2008) trained fifteen college students to establish two six-member equivalence classes using a standard MTS simultaneous training protocol and a feed back fading procedure (e.g. the feedback is provided 100% of trials, after performance were stabilized, feedback were reduced to 50% of trials, finally feedback dropped to 0%). In the transfer of function training, each participant was trained to press J key seven times in the presence of stimulus C1, and press J key three times in the presence of stimulus C2 in a randomized block with feedback fading procedure, then tested for the single-response transfer amongst equivalence members. New responses were then trained to C and D stimuli, followed with the test of dual-option transfer amongst equivalence members. Only four participants demonstrated equivalence performance despite intensive training. No nodality was found in correct responses. The proportion or frequency of responses transferred to other members of equivalence class in terms of responses trained to C stimuli in single-option transfer test confirms the assumption of interchangeability and substitutability amongst equivalence members, but not provide evidence of a structural account for equivalence formation. However, interestingly, 3/4 participants who formed equivalence demonstrated dual responses transfer to other equivalence members consistent with discriminative responding trained to C and D stimuli. That is, proportion or frequency of response transferred from C stimuli was larger to A and B stimuli, and not to E and F stimuli; proportion or frequency of response transferred from D stimuli was larger to E and F stimuli, and not to A and B stimuli. The authors suggested that this bifurcation of response transfer can not be attributed to other variables such as differential reinforcement, unequal trial presentation (Imam 2001, 2006), rather the format of testing emergent stimulus relation might be a plausible explanation of why nodal effect was observed in some studies, but not others. They further argued that there is "coexistence of the interchangeability of stimuli in an equivalence class and the bifurcation of such a class in terms of nodal structure" (p. 359) as a theoretical adaptation to Sidman's account (1994, 2000). That is, the reinforcement contingency established by conditional discrimination training not only create class-specified discrimination among the stimuli in different classes, but also node-based discrimination among the stimuli in the same equivalence class. The expression or non expression of either the two discriminative functions was determined by the different formats used in testing the emergent relations (see also Fields & Moss, 2007 for a systematic review).

In addition to this, a number of animal researchers have argued that nodal distance is established by conditional discriminations resulting in differential reinforcement history amongst stimuli (Treichler & Van Tilburg, 1996; Bond et al., 2003; Vasconcelos, 2008). In the animal literature, nodal distance is referred to as the *symbolic distance effect* (SDE), in which, both human and animal's response accuracy increases and reaction time decreases as the number of intervening terms between the two test stimuli increases (Bryant and Trabasso, 1971; Bond et al., 2003; Wynne, 1997; D'Amato & Colombo, 1990). In contrast to the nodal number effect involved in the human equivalence literature, response accuracy was reported to decrease and reaction time to increase as the number of nodes increased (Bentall, et al, 1993; 1998; Dube, et al., 1993; Holth & Arntzen, 2000; Lazar, 1977; Wulfert & Hayes, 1988; Spencer & Chase, 1996; McDonagh, Mcllvane, & Stoddard, 1984; Fields et al., 1990; Dymond & Refeldt 2005). However, one familiar with animal research would realise the two assumptions indeed involve different processes, despite the same terminology being used. The symbolic distance effect can be accounted for as the further apart two stimuli are in an ordered series, the faster/more accurate a participant is in responding to which *stimulus* is greater or lesser, hence the participant has to make a comparative judgement between stimuli in a given pair, as described in Section 1.3.2.2. It seems that the ability to respond discriminately between stimuli in animals and the so called "transitive inference" responses are far away from the relational responding commonly observed in humans. This cautionary note clarifies the confusion caused by the similar terminology used in the two fields of research. In order to further distinguish our account "nodal distance", "nodal number effect" or "nodality" is used throughout the current thesis.

In summary, Section 1.3 outlined the ongoing debate in the equivalence literature from human to animal studies, from conceptual development to methodological verification. However, a number of questions remain unanswered, such as: Are there other variables, apart from reinforcement, that also account for equivalence? Are the members of an equivalence class interchangeable or substitutable with each other? And more importantly, is equivalence the basic unit in language acquisition and reasoning? These questions will be address empirically in Chapters 2 and 3 with the systematic manipulation of reinforcement, and discrimination.

1.4 Stimulus Control from a Cognitive Perspective

The importance of the role of reinforcement in equivalence formation (Sidman 1994, 2000) or the formation of derived stimulus relations (Hayes, et al, 2001) also suggests a potential role in language acquisition, as equivalence and derived stimulus relations are regarded as critical to understanding the semantic processes of language acquisition (Wulfert & Hayes, 1988; Hayes & Hayes, 1992; Catania, 1998; Dugdale & Lowe, 1990). On the other hand, the cognitive tradition in the area has also provided some interesting findings to account for this complex issue. Two features known as Age of Acquisition/ Time of Acquisition (AoA) and word frequency were thought to prove some evidence on stimulus control in semantic generalisation will be presented next.

1.4.1 Age of Acquisition (AoA)

It is often the case that words learned earlier are also recognised, named, and categorized faster than words learned later, this phenomenon has been widely observed and is known as "Age of Acquisition" (AoA) (Brysbaert, 1996; Brysbaert, Lange, et al., 2000; Brysbaert et al., 2000; Gerhand & Barry, 1999; Morrison & Ellis, 1995, 2000; Zevin & Seidenberg, 2002; Izura and Ellis, 2002). A series of experiments investigating AoA in dominant Spanish-English bilinguals (Spanish native) with controlled frequency, word length, and object familiarity were conducted by Izura and Ellis (2002). In Experiment 1, 32 dominant Spanish-English bilinguals were equally assigned to two experimental groups, half of them were asked to name the object presented successively on the screen in English, whereas the other half had to name the object array in Spanish. Analysis of variance (ANOVA) in response time (RT) of correct responses demonstrated faster responses in earlier acquired objects then later acquired objects in both Spanish and English naming. Greater errors in later compared to earlier acquired English words were found using Wilcoxon nonparametric tests. Experiment 2 aimed to replicate the findings of Experiment 1 in a word recognition task (the lexical decision task). 22 dominant Spanish-English bilinguals were randomly assigned to two groups, one used English words and nonsense words successively and the other group used Spanish words and nonsense words successively. Each participant was instructed to press the P key on the keyboard if the stimulus was a word and to press the Q key if it was a nonsense word. RT and error data replicated the AoA observed in Experiment 1. These findings supported the prediction that it is the order and not the age at which words are acquired that is responsible for the AoA effect (Davis & Kelly, 1997).

1.4.2 Frequency

Another well documented phenomenon in language acquisition is known as the frequency effect, that is, words that are encountered more often (high-frequency) are learned faster than words that are encountered less (low-frequency) (Brysbaert, 1996; Brysbaert et al., 2000; Brysbaert, Lange, et al., 2000; Gerhand & Barry, 1998, 1999). For example, Brysbaert, Lange and Wijnendaele, (2000, Exp1) used six lists of 24 Dutch words in testing immediate naming, delayed naming, and a lexical decision task, respectively. The first pair of lists differed in frequency and was matched on AoA and imageability. The second pair of lists differed in AoA and was matched on frequency and imageability. The third pair differed in imageability and was matched on frequency and AoA. 60 university students demonstrated that RT of correct responses were faster in Dutch words had learned earlier and encountered more often than Dutch words had learned late and encountered less in both naming and lexical decision tasks. However, no effect of imageability was found when AoA and frequency were controlled.

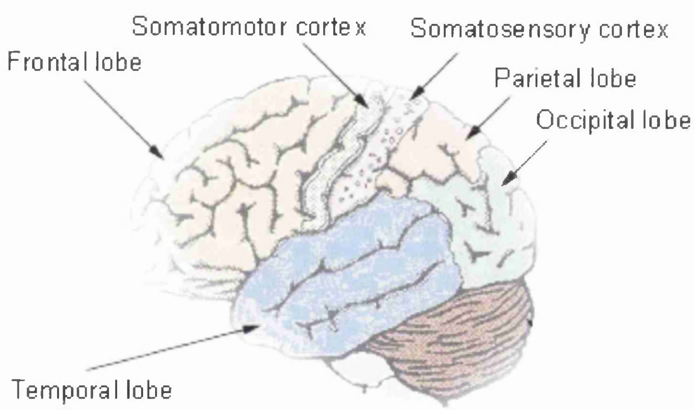
Recently, there is growing evidence that the effects of AoA and frequency are highly correlated in a variety of word processing tasks, such as the experiment described above. Because many high-frequency words are acquired relatively early in life and many low-frequency words are acquired relatively late in life (Ghyselinck et al., 2004). This has led to the prediction that the strength of a new word added to the person's existing language repertoire (e.g. learned, recognized) is a cumulative effect from both AoA and frequency, hence, the weight of both effects in word processing is the same. However, this assumption is challenged by traditional connectionist models, that emphasis the role of AoA over frequency, and the growing network model (Steyver & Tenenbaum, 2005), which gives no specification of the weight between AoA and frequency. (see Ghyselinck et al., 2004 for a review).

In summary, Section 1.4 provided a cognitive account of stimulus control in semantic generalisation, which suggested that the time when a word is learned and how frequently it is encountered, determines the strength of association of a word added into an existing semantic network. Although there is evidence for a correlation between AoA, and frequency, no empirical study has attempted to address whether there is any interaction between AoA, frequency and reinforcement. Chapter 4 of the current thesis aims to address this issue.

1.4 Neurological basis of Equivalence Class

There is increasing evidence that equivalence class members produce the same neurological activity as semantically related words in the priming literature (Barnes-Holmes, et al, 2005; Yorio, et al, 2008; Haimson, et al, 2009), some basic neurological background on language acquisition will be summarised below. The connection between language and brain function has derived from the study of localization of brain function by neurologists and neurosurgeons, which originated from studies with Brain Injured Patients who had suffered brain damage in certain regions of cortex that resulted in some form of language disability (e.g. the discovery of Broca and Wernicke's areas, see Purves, et al., 2008). Based on brain anatomy, several brain regions have been implicated in language development. Specifically, Broca's area located in the ventral posterior region of the frontal lobe in the left hemisphere has been linked to speech production (i.e., sentence structure and utterance). Wernicke's area located in the posterior and superior regions of left temporal lobe to language comprehension (i.e., understanding the meaning of words). The primary visual cortex located on the occipital lobe plays a role in visual perception (i.e., word perception). The primary motor cortex located on the somatomotor cortex is linked to motor control (i.e. the movement of the tongue), the primary somatosensory cortex located on the sematosensory cortex is linked to sensory information and the primary auditory cortex located in superior region of the temporal lobe is linked to auditory information (i.e. hearing, pitch, volume). Complex interactions of these regions are thought to provide the physiological basis for language. See Illustration 1 for the different lobes in the brain.

Illustration 1: The lobes in the brain (cerebrum)

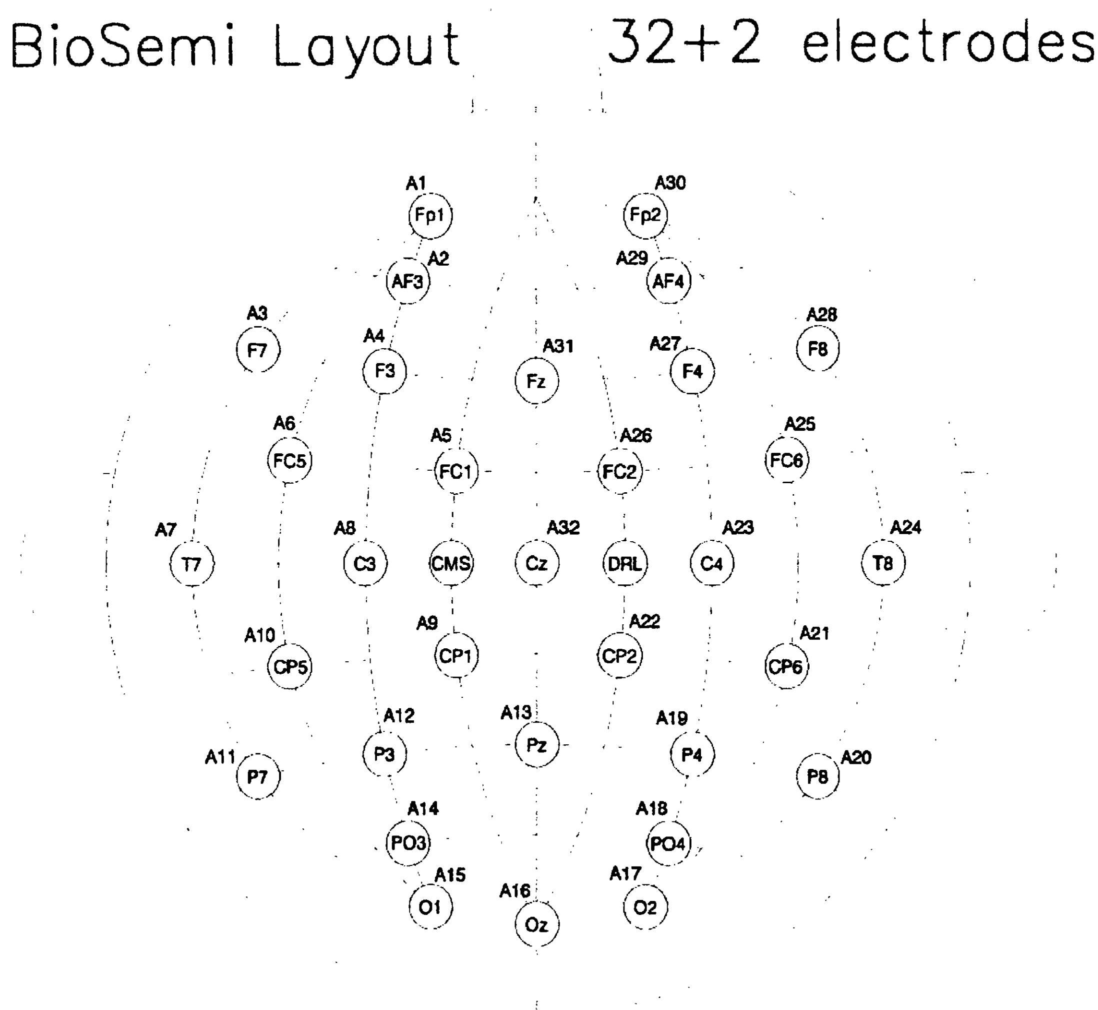


Lobes of the cerebrum

The brain-language link has primarily been demonstrated in the left hemisphere (e.g. Broca and Wernicke's areas), but research in the area of language function has not been exclusive to the left hemisphere, rather the lateralization of hemispheric function of language may reflect simply dedication of both hemispheres to significantly different but complementary functions (Taylor & Regard, 2003). This assumption is not only demonstrated from studies of electroencephalography (EEG) in semantic processing (Kutas & Hillyard, 1980), but also in studies of Positron Emission Tomography (PET) (Damasio, et al., 1996), and Functional Magnetic Resonance Imaging (fMRI) (Martin, et al 2000) in categorization. The latter two are both functional neuroimaging techniques, which "light up" brain regions when activated during cognitive processing, and allow us to "see" the internal brain functions. The current thesis, relies on the former technique (EEG), therefore more information about functional neuroimaging is beyond the current research scope, interested readers can read a systematic behavioural review of functional nuroimaging and derived stimulus relations by Dickins (2005).

EEG was developed by neurophysiologist Hans Berger in 1929. In his seminal work he showed that one could measure the electrical activity of the human brain by placing an electrode on the scalp, amplifying the signal, and plotting the changes in voltage over time on a histogram. The source of this electrical activity comes from the changes in iron permeability that alters the distribution of electrical charges across the neuronal membrane, thus changing the membrane potential of the affected neurons (Purves, et al., 2008, Principles of Cognitive Neuroscience, p.13). Modern EEG recording uses a set of electrodes that are typically embedded in an elastic cap and applied to the scalp (see Illustration 2 for a scalp location of a typical set of 32-electrodes plus the Common Mode Sense active electrode (CMS) and the Driven Right Leg passive electrode (DRL).

Illustration 2: Scalp location of a typical set of 32 + 2 electrodes



EEG signals are analyzed in terms of the power in various frequency bands at each electrode location, the major bands of interest being delta (up to 4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-20 Hz), and gamma (30-70 Hz) rhythm. See Table 1 for a summary of functions commonly associated with these brain rhythms.

Table 1: A summary of functions commonly associated with major brain rhythms.

Type	Frequency(Hz)	Functions
Delta	Up to 4	Slow wave sleep in adults; commonly in babies; in continues attention tasks (Kirmizi-Alsan et. al. 2006)
Theta	4-8	Drowsiness or arousal in adults, meditative status (Cahn & Polich, 2006)
Alpha	8-12	Closing the eyes, relaxing (Berger, 1929)
Beta	12-20	Active concentration (Pfurtscheller & Lopes da Silva,1999)
Gamma	30-70	Cross-model sensory processing (Kisley & Cornwell 2006)

The use of EEG has several advantages over behavioural and other physiological measures. First, it provided the first opportunity to study brain electrical activities with a noninvasive approach so that surgical procedures were no longer required in neurological studies. Second, EEG can be used when there is no behavioural response required, and provides vital information on neurological processes underneath skull. Third, EEG can detect changes on a millisecond timeframe, whilst, PET and MRI have time resolution from seconds to minutes, thus allowing a more accurate time link between brain activity and behaviour. Forth, EEG directly measures the brain's electrical activities, whilst, PET and fMRI indirectly measure the brain's electrical activities through changes in metabolic activity (PET) and blood flow (fMRI), which can be distorted with other variables, such as abnormality in metabolic activity or blood flow. Finally, it is a relatively cost effective technique compared to PET or fMRI.

However, there are several limitations to EEG. First, the psychological function behind EEG signals is never as clear as those measures from behavioural data (Luck, 2005). The ongoing EEG record reflects the summed activity of all ongoing processes in the brain region monitored by the electrodes, thus is way too broad to relate to specific cognitive functions. Second, EEG has poor spatial resolution compared to PET and fMRI. Third, EEG only measures electrical activities from open fields in the brain but not closed ones (important components situated deep inside the brain, e.g. hypothalamus).

A more optimal solution for relating EEG signals to cognitive functions has been developed by extracting and averaging time-locked (events of interest only) EEG segments to generate an averaged (grand) histogram across participants, known as "Event-Related Potentials" (ERPs) (Kutas & Hillyard, 1980). One potential weakness with ERPs, is that their analysis requires a large amount of data per condition, thus, limiting its power. Despite the need for larger data sets ERPs is the most commonly employed method for investigating direct brain electrical activities during language based tasks. The first study to employ this technique in the area of language function was that of Kutas and Hillyard (1980). In their study, participants' on going EEG was recording during a sentence reading task. The sentences either ended normally, or were completed by unexpected words that either involved a semantic violation, or a physical violation or both. An enhanced negative waveform was observed 400 milliseconds (N400) after the presentation of semantically unexpected word at the end of a sentence (e.g. Emma wears *dog*) compared to a semantically appropriate word at the end of a sentence (e.g. Emma wears *shoes*). The N400 to semantic violations/mismatch in priming effect has been widely demonstrated in the literature ((Kutas & Hillyard, 1980, 1983, 1984; Anderson & Holcomb 1995; Deacon, et al 2004, Hinojosa, Martı´n-Loeches, & Rubia, 2001; Heinze, Muente & Kutas, 1998).Whereas, an inflated positivity observed 300 milliseconds (P300) after the presentation of physically aberrant word at the end of a sentence (e.g. Emma wears *SHOES*) compare to physically normal word at the end of a sentence. Therefore, if the N400 effect was observed between non equivalent and equivalent relations, the assumption that equivalence or derived stimulus relations provided a behavioural account of semantic network formation would be further verified (see Chapter 5 of the current thesis for an experimental investigation of this effect).

Logically, semantic associations in language learning also necessitate semantic categorization during learning. Azizian and his colleagues (2006) argued that P300 might serve as a neural marker for perceptual categorization. In their target recognition task, human faces with different number of distinct features (e.g. one nose, two eyes) were categorized into nine groups. Their findings suggested that the P300 was larger when non-target stimuli were perceptually similar to the target stimulus than it was for other non-target stimuli. Other studies in cognitive neuroscience suggested that the P300 is larger when subjects devote more effort to a task, leading to the postulate that the P300 amplitude can be used as a measure of selective attention and source allocation (Donchin & Coles, 1988; Isreal et al., 1980; Johnson, 1988; Gray, et al., 2004). If these assumptions are valid, it would be predicted that a larger P300 in stimulus equivalence relations would emerge than for non equivalence relations, and this would be more apparent for nodal relations that were separated by more nodes.

In summary, EEG measures have provided the opportunity to directly examine neurological change during behavioural tasks. If stimuli participating in an equivalence class produce the same neurological changes as members of a semantic network in the priming literature, then the assumption that stimulus equivalence provides a behavioural account of language acquisition will be further supported. Chapter 5 of the current thesis aims to address this issue.

1.6 Aims of the current Thesis

The current thesis draws on behavioural, cognitive, and neuropsychological evidence, in order to provide a multi-disciplinary account of equivalence class formation as a model of language acquisition. The first two empirical chapters aim to provide a systematic investigation of the delayed emergence of equivalence using a MTS paradigm. Chapter 2 comprises four experiments that systematically manipulate training structure in equivalence class formation to examine whether nodal number effects are the result of unequal reinforcement history during conditional discrimination training, while the last two experiments replicate Fields and Watanable-Rose's (2008) finding that the nodal number effect is maintained in the testing of transfer of function. Chapter 3 provides the first empirical test of Saunders and Green's (1999) conditional discrimination hypothesis in explaining nodal number effects in equivalence class formation. The last two empirical chapters aim to bridge the gap between behavioural, cognitive and neuropsychological research on language acquisition. Specifically, Chapter 4 attempts to systematically manipulate reinforcement, age of acquisition and word frequency in investigating semantic network growth (i.e., equivalence class formation). Chapter 5 incorporates an additional dependent measure (EEG) while testing the effect of levels of reinforcement on nodal distance in equivalence class formation. Together the empirical work reported herein aims to add to the literature on equivalence class formation as a model of language acquisition.

Chapter 2: Preserved Nodal Number Effects under Equal Reinforcement

2.1 General Introduction

Several laboratory studies have demonstrated that when a number of interrelated conditional discriminations are trained, *derived* (untaught), relations often emerge, even though the stimuli do not necessarily share any physical properties in common with one another (Sidman, 1971). Typically, in these studies, the minimum number of interrelated conditional discriminations are trained and then the derived relations are tested (e.g., train A→B and B→C, then test C→A). The term *training structure* has been used to refer to the sequence of these conditional discriminations and the arrangements of linking stimuli presented in baseline training. A *node* has been defined as a stimulus that is linked by training to at least two other stimuli (Fields & Verhave, 1987; Fields, Verhave, & Fath, 1984) and the number of nodes that link any two stimuli in a set of trained conditional relations is described as the *nodal number* (Sidman 1994). For example, a 5-member class (A, B, C, D, and E) contains six 1-node relations (e.g., B-D, with C as the node), four 2-node relations (e.g., B-E, with C and D as nodes), and two 3-node relations (e.g., A-E, with B, C, and D as nodes).

The predicted test performance for derived relations often does not emerge immediately and usually requires repeated exposures to the training and testing phases (Lazar, Davis-Lang, & Sanchez, 1984; Sidman, Willson-Morris, & Kirk, 1986). This is known as *delayed emergence*, and is reflected in the higher probability of incorrect responses on earlier, rather than on later, trials. Several studies have also reported that the probability of successful emergence is a function of nodal number (Bentall, Jones, & Dickens, 1998; Dube, Green and Serna, 1993; Fields, Adams, Verhave, & Newman, 1990; Kennedy, 1991; Lazar, 1977; Dymond & Rehfeldt 2005; Spencer & Chase, 1996; Wulfert & Hayes, 1988). Fields et al. (1990), for example, demonstrated an interaction between these factors: one-node relations initially exerted more control than the 2-node relations during testing for equivalence, and eventually all relations exerted complete control.

Accuracy of responding in unreinforced probes for derived relations has been the most common measure of the relatedness of stimuli. In addition, supplemental measures of derived relational responding may shed more light on the nature of the relations among stimuli (Dymond & Rehfeldt, 2001). Several researchers have reported that response speed were a function of nodal number (Bentall, Dickens, and Fox, 1993; Bentall et al., 1998; Holth & Arntzen, 2000; Spencer & Chase, 1996; Wulfert & Hayes, 1988) even when accuracy remained intact. Fields, Landon-Jimenez, Buffington and Adams (1995; see also Fields, Adams, Verhave, & Newman, 1993) adopted another alternative measure of stimulus relatedness. In this study, two 5-member equivalence classes were trained using a protocol that ensured equal reinforcement across trial types. All 12 participants passed baseline discriminations, although only 2 formed equivalence classes. After these two participants demonstrated the formation of equivalence classes, new responses (i.e. pressing the J key 3, 5, 7, 9 times, respectively) were trained to the end, A and E, stimuli in each group. Response transfer was measured by the relative frequency with which responses trained to A and E was evoked by all stimuli in both classes. In general, response transfer was an inverse function of nodal number.

The status of nodal number as an independent variable has been questioned, however, and it has been suggested that apparent nodal number effects are a function of other variables (Imam, 2001, 2003, 2006; Sidman 1994, 2000). Imam has noted that Sidman's account of equivalence does not include the notion that test outcomes should vary as a function of training structure, order, or direction, providing that extraneous stimulus control is prevented (Carrigan & Sidman, 1992; Sidman, 1994). Sidman (2000) has suggested that "procedural factors that might account for the results of experiments that have given rise to notions of directionality and nodal distance" (p. 145).

In Imam (2001; Experiment 1), three 5-member equivalence classes were trained in a serial manner, across two conditions: accuracy only and accuracy with a *limited hold* (LH; the participant was only given positive feedback to correct choices that occurred within a specific time period). The serial training procedure involved training the AB relations to criterion, then introducing the BC relations (i.e., mixed AB and BC), and so on. In this way, the number of reinforcers scheduled for responses to particular trial types was deliberately unequal. Response latency in Experiment1 tended to be an inverse function of nodal number on transitivity trials, but only in accuracy-only condition on equivalence trials. That is, a nodal number effect was generally observed, although a time-accuracy trade off seemed to have occurred. In Experiment 2, a protocol ensuring equal numbers of reinforcers were scheduled across trial types was implemented and the class size was increased to seven members. The participants' response time in Experiment 2 tended *not* to decline as the nodal number between stimuli increased. According to Imam, "By equalizing reinforcement history, the confound noted in the first experiment was eliminated, and the nodal number effect observed in the second experiment thus was greatly diminished for one- through five-node trials" (2006, p. 109).

Imam (2003) replicated Imam's 2001 Experiment 2 in a single participant. The participant formed four independent three, 7-member equivalence classes and showed transfer of time under the two transfer conditions used. Response time was not a function of nodal number, repeating the effect seen in Imam (2001, Experiment 2) when the number of reinforcers was equal across trial types. Imam (2006) established different sets of three 7-member equivalence classes across 4 participants by using a within-subject comparison of simple-to-complex, complex-to-simple, and simultaneous protocols. The protocols were implemented under either accuracy-only or accuracy-plus-time conditions while keeping number of presentations of training and testing trials equal. Again, response time and accuracy did not decrease as a function of nodal number, with or without the time contingency, or under any protocol.

According to Imam's interpretation of his studies, it appears that nodal number effects reported in previous studies may have been the product of a procedural artifact, namely unequal reinforcement during baseline discrimination training. There were, however, a number of problems with the statistics employed, and interpretation of those statistics, in Imam (2001, 2006). A repeated-measures analysis of variance (ANOVA) was employed to assess differences in RTs as a function of nodal number for each participant. It is incorrect, however, to conduct a repeated-measures ANOVA using the same participant's data more than once in each condition because the influence of the particular participant is considered in relation to the population mean. Therefore, a key assumption of a within-participant ANOVA is that participant effects are "assumed to be *independent* of each other" (see Girden, 1992, p.8).

Irrespective of this issue, other factors were likely contributors to the significant effect in Experiment 2.1 and the *lack* of a significant effect in Experiment 2 (Imam, 2001, see also Imam, 2003, 2006). In Experiment 2.1, three conditions (1-, 2-, and 3-node) were compared using an ANOVA. In Experiment 2, however, five conditions (1-, 2-, 3-, 4, and 5-node) were compared. However, a comparison of the effect size (η^2) in Imam (2001) shows that nodal number accounted for *more* of the total variance in Experiment 2 than in Experiment 2.1 (Table 9 vs. Table 5, respectively). It is impossible to conclude, therefore, that nodal number effects disappeared in Experiment 2, a conclusion that can *only* be reached when power is high (Loftus, 1996), in the absence of a quantifiable hypothesis. Similarly, Imam 2003 and 2006 rejected the null hypothesis (H^0) when nodal numbers were compared across 5 conditions, but again it is impossible to determine if this is because the 5-node study had less power than the 3-node study, or because the manipulations had an important effect. Parenthetically, it is incorrect to maintain that a H^0 is "true", although it can be rejected "for all intents and purposes" (Loftus, 1996, p. 164) *if* power is high, and means are roughly equal, which was not the case in the Imam studies.

A potential confounding factor in the Imam (2001) study was that training structure was varied across Experiments 1 and 2. In Experiment 1, AB relations were trained to criterion, followed by a mix of AB and BC trial types, and so on. In Experiment 2 all trial types were introduced simultaneously (i.e., AB, BC, DE etc. were all presented in a mixed training block from the beginning). It is possible that differential training protocol *per se* was responsible for the elimination of nodal number effects in Imam (2001) Experiment 2. Although the rationale behind the Imam (2001) was to test the role of reinforcement history and not the role of training structure, the key point of that study was that the effects of nodal number could be eliminated by balancing the number of training trials (and presumably also therefore the number of delivered reinforcers).

There were also a number of other methodological differences between the Imam studies and other studies of nodal number that may have contributed to the failure find a nodal number effect. Imam (2001) presented three stimuli on the screen at the same time, although Kennedy (1991) found that the addition of a third stimulus had the effect of attenuating nodal number effects. The use of a within-group design when employing accuracy and RT as measures of equivalence learning is also questionable. Response patterns of non-naïve participants have been shown to be different to those of naïve participants during equivalence tasks (Fields et al., 1997): in this case, a between group design is obligatory in order to avoid either a main effect of previous equivalence class formation or an interaction of class formation with condition (see Greenwald, 1976).

The findings of the Imam (2001, 2003, 2006) studies are particularly important because they suggest that previous reports of nodal effects were due to a methodological confound: we have outlined some potential problems with the Imam studies, however. Therefore, the aim of the present set of research was to manipulate reinforcement and training structure, and then examine if nodal effects would indeed disappear under equal reinforcement. A group design was employed (Fields et al., 1997) in Experiment 2.1 and 2.3. In the *Unequal Reinforcement* condition, each trial type was introduced serially in training, whereas in the *Equal Reinforcement* condition, all trial types were randomly intermixed from the beginning (i.e., a simultaneous training protocol). The equal reinforcement without limited hold condition was replicated in Experiment 2.2 with a larger sample, and in 2.4 with increased baseline training criterion. Nodal number was measured as a function of response accuracy, (Experiments 2.1, 2.2, 2.3, and 2.4), response speed (Experiment 2.1 and 2.2), and in a response-transfer test (Experiment 2.3 and 2.4) in two 5-member (Experiment 2.1 and 2.2) and two 6-member (Experiment 2.3 and 2.4) equivalence classes. In Experiment 2.1, If nodal number effects are indeed a function of a particular training protocol and/or differential reinforcement history, then differential response accuracy, response speed and transfer of function should only be observed in the Unequal Reinforcement condition. Experiment 2.2 will replicate the equal reinforcement without limited hold condition with an increased number of baseline training trials. In Experiment 2.3 and 2.4, a new version of the *transfer paradigm* will also be tested to assess the nodal effects in post-equivalence formation.

2.2 Experiment 1

Experiment 1 sought to examine whether nodal effects are a result of unequal reinforcement during conditional discrimination training. A LH was also employed, in order to examine if a time-accuracy trade off would occur, and thereby mask any nodal effects (Dickens, 2005). If nodal effects were apparent under no LH, but disappeared or were weakened under a LH, then this might explain the results of previous research that failed to report nodal number effects when a LH was in place (Imam, 2001, 2003).

2.2.1 Method

2.2.1.1 Participants

Twenty-two adults participated in Experiment 2.1 (11 male; 11 female), ranging in age from 18 to 62 years (mean age = 27.26 years, standard deviation = 12.05 years). Nineteen participants were students (11 undergraduate, 8 postgraduate) at Swansea University, one participant was retired, and the remaining participant was a Human Resources Officer. Participants were recruited through email and word of mouth from the experimenter and all were naïve about the purpose of the experiment. In return for their participation, participants earned £3 for each session, which was not contingent upon performance. The study was approved by the Department of Psychology, Swansea University, Ethics Committee.

2.2.1.2 Apparatus and Materials

The study was conducted in a quiet room, containing only a desk, a chair and a Personal Computer with a 550 MHz processor, a 14-inch color monitor, and a standard computer mouse. Each participant sat at a table facing the computer monitor and keyboard. The computer controlled all trial presentations, trial and phase order, and recorded all responses and RSs.

The stimuli were obtained from Massaro, Venezky and Taylor (1979) and were letter permutations derived from the most frequent 150 six-letter English words as listed in Kucera and Francis (1967). These pseudoword stimuli (Whelan et al, 2005) met the following criteria with reference to the English language: (i) they were orthographically regular; (ii) they were pronounceable; (iii) they contained common vowel and consonant spellings, and (iv) they had no more than three letters for a medial consonant cluster, if one occurred (i.e. boceem, lewoly, matser, and lorald; see Table 2). The assignment of stimuli was randomized across participants. The stimuli were in black Times New Roman font, set against a white background.

Table 2: Stimuli employed in all four experiments (Assignment of stimuli was randomized across subjects).

boceem	vartle	rettes	drager	siflet	troper
lewoly	lorald	rigund	surtle	gedeer	haveen
matser	betret	copher	casors	wollef	ronkeb
samolt	desund	cachen	murben		

2.2.1.3 Procedure

A short questionnaire was administered to record participants' age, gender, occupation and previous knowledge of the research topic and each participant was also given a consent form to read and sign before beginning the study. All participants were exposed individually to the experimental procedure across four sessions (defined below), irrespective of performance in the experimental task. These sessions were scheduled over a 1-week period and each lasted between 20-35 mins.

Each participant was randomly assigned to one of the four conditions, which differed in terms of training structure and presence or absence of an LH, and are described using the following nomenclature: Unequal Reinforcement no LH; Unequal Reinforcement LH; Equal Reinforcement no LH, and Equal Reinforcement LH.

At the start of the experiment, the following instructions appeared on the computer monitor:

> "In a moment a word will briefly appear in the middle of the screen. It will disappear and two other words will appear. Choose 1 of the 2 words in the corner of the screen by pressing the Z key for the left word and the M key for the right word. During some stages of the experiment, the computer will NOT tell you if your choices are correct or wrong. However, based on what you have learned so far, you can get all of the tasks correct. Please do your best to get everything right. Thank you and good luck."

For the two experimental conditions that included an LH (participants had to respond within 2.5 s), instructions also included the phrase "It is important that you respond as quickly as possible!"

Each trial started with a 1.7-s presentation of a sample stimulus, at the centre of the screen, which disappeared and was replaced by the two comparison stimuli that appeared after a 1-s interval. Participants pressed the 'Z' or 'M' key on the computer keyboard to select the comparison on the left or right respectively. When feedback was provided, choosing the correct comparison produced a 1-s display of the word 'Correct'. Choosing the incorrect comparison produced a 1-s display of the word 'Wrong'. Both were displayed in brown in the middle of the computer screen, and were followed by a 1.5-s intertrial interval (ITI), during which the screen was blank. In the LH conditions, if participants failed to choose one of the comparison stimuli within 2.5 s, the phrase "Timed Out" appeared in maroon at the top of the computer screen.

Two 5-member equivalence classes were established by training AB, BC, CD and DE relations (i.e., a linear structure). All trial types were presented randomly within a block. All types of training block were followed by informative feedback on the participant's choice of comparison. In the Unequal Reinforcement no LH and Unequal Reinforcement LH conditions, the AB trials were first trained to the mastery criterion of 8 consecutively correct responses. Next, mixed AB and BC trials were presented until the same mastery criterion was reached, whereupon a new trial type was introduced, and so on until all trial types were presented in a mixed block. Eight consecutively correct responses were required on the final mixed block to proceed to the test phase. In the Equal Reinforcement no LH and Equal Reinforcement LH conditions, the two equivalence classes were established by training AB, BC, CD and DE relations on a simultaneous basis, in order that each relation was presented the same number of times. That is, these trial types were presented in a random manner in a mixed block from the beginning of the experiment. Eight consecutively correct responses were required to proceed to the test phase.

Once the criterion for the training session had been met, the test phase commenced without warning and the corrective feedback (i.e., not including the "Timed out" feedback) terminated. All baseline conditional relations, tests for mutual entailment, 1-, 2- and 3- node transitivity and equivalence trials were presented in a single randomized block. Each type of relation was presented the same number of times with 40 trials in total (see Table 3).

Table 3 Trial types per relation type that were presented in Experiment 1.

Relation Type	Trial Type					
Directly Trained						
	A1B1	B1C1	C1D1	D1E1		
	A2B2	B2C2	C2D2	D2E2		
Symmetry						
	B1A1	C1B1	D1C1	E1D1		
	B2A2	C2B2	D2C2	E2D2		
1 Node						
	A1C1	B1D1	C1E1	C1A1	D1B1	E1C1
	A2C2	B2D2	C2E2	C2A2	D2B2	E2C2
2 Node						
	A1D1	B1E1	D1A1	B1E1		
	A2D2	B2E2	D2A2	B2E2		
3 Node						
	A1E1	E1A1				
	A2E2	E2A2				

A *cycle* was defined as training all relations to criterion and testing all possible derived relations. Participants were exposed to 2 cycles in each session across a total of 4 sessions in a 1-week period, regardless of their performance.

2.2.2 Results

Twenty-two participants began Experiment 1, 5 participants in Unequal Reinforcement no LH (4 passers), 4 participants in Unequal Reinforcement LH (4 passers), 8 participants in Equal Reinforcement no LH (4 passers), and 5 participants in Equal Reinforcement LH (3 passers).

The actual number of reinforcers delivered in Experiment1 across all participants is presented in Table 4. These data indicate that the procedures employed were successful in manipulating the number of reinforcers across trial types for most participants in both conditions, with minor variations. In the Unequal LH and no LH conditions, all participants (pt1, 2, 4, 5, 6, 7, 8, and 22) except Participant 3, the most number of reinforcers were delivered during AB trials, the next highest amount during BC trials, then CD trials, and the lowest number of reinforcers during DE trials. The number of delivered reinforcers was higher for DE trials than for CD trials for Participant 3. In contrast, the number of delivered reinforcers was approximately equal across all trial types for all participants (pt 9, 11, 12, 14, 16, 17, 18, 19, 20, and 21) except Participant 10, 13, 15, in the Equal Reinforcement LH and no LH conditions. More reinforcers were delivered for BC and CD trials in Participant 10 and 13, while, Participant 15 received more reinforcers on AB and BC trials.

Table 4 Number of reinforcers delivered per training block per participant

Block	Trial Type															
	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE
	Participant 1				Participant 2				Participant 3				Participant 4			
1	28	22	13	4	25	13	7	4	45	20	20	4	27	10	4	2
2	17	12	8	3	24	14	4	2	23	8	5	14	28	14	11	2
3	18	8	4	2	49	26	15	8	18	13	3	14	22	13	10	1
4	19	11	7	2	42	26	7	5	23	12	7	20	27	17	5	2
5	16	9	5	2	17	8	5	2	28	18	7	13	17	8	5	2
6	17	8	5	2	17	9	6	2	16	8	6	9	17	8	5	2
7	21	13	4	3	18	8	4	2	16	9	4	9	17	8	5	2
8	16	10	5	3	17	8	5	2	14	5	4	1	15	10	5	2
Total	152	93	51	21	209	112	53	27	183	93	56	84	170	88	50	15
	Participant 5				Participant 6				Participant 7				Participant 8			
1	66	22	8	5	52	42	11	5	35	19	14	4	32	15	15	5
2	27	21	10	8	35	16	13	7	17	12	5	2	29	16	8	3
3	20	11	7	3	29	18	9	2	22	9	5	2	18	8	4	2
4	28	9	6	5	19	12	9	5	15	9	6	2	16	9	5	2
5	17	9	4	2	25	14	8	2	16	8	6	2	18	9	8	4
6	16	9	5	2	20	10	8	3	17	8	5	2	16	8	6	2
7	17	9	4	2	18	11	7	2	16	8	6	2	18	11	7	2
8	16	9	5	2	16	9	6	1	16	8	6	2	18	9	6	3
Total	207	99	49	29	214	132	71	27	138	73	47	16	165	85	59	23
	Participant 9				Participant 10				Participant 11				Participant 12			
1	12	16	12	2	29	28	35	21	9	5	7	7	8	5	7	6
2	2	3	4	2	10	10	11	7	2	2	2	2	1	6	6	6
3	2	2	2	2	11	23	20	11	2	2	2	2	3	4	3	5
4	2	2	2	2	5	16	13	15	2	2	2	2	1	2	4	1
5	2	2	2	2	2	3	2	4	2	2	2	2	2	2	2	2
6	2	2	2	2	5	11	8	10	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	2	2	6	5	6	4	2	2	2	2
8	2	2	1	0	2	2	2	2	2	1	2	3	2	2	2	2
Total	26	31	27	14	66	95	93	72	27	21	25	24	21	25	28	26

	Participant 13				Participant 14				Participant 15				Participant 16			
1	17	21	19	11	10	9	9	8	29	26	14	19	2	3	3	3
2	6	8	8	8	2	3	2	1	12	9	5	4	20	22	22	23
3	2	4	6	3	2	2	2	2	3	4	3	4	6	7	7	6
4	1	3	2	2	2	2	2	2	2	2	1	3	1	2	3	2
5	4	4	5	3	2	2	2	2	4	4	2	3	8	7	7	8
6	1	3	1	3	2	2	2	2	1	2	3	2	14	13	12	15
7	2	2	2	2	3	4	2	2	2	2	2	2	2	2	2	2
8	3	4	3	2	2	2	2	2	2	2	2	2	3	2	2	2
Total	36	49	46	34	25	26	23	21	55	51	32	39	56	58	58	61

	Participant 17				Participant 18				Participant 19				Participant 20			
1	15	15	15	14	20	20	19	20	18	18	18	18	7	8	8	7
2	10	11	11	9	12	11	10	11	5	5	5	5	9	6	6	7
3	8	8	9	9	16	16	14	16	3	2	4	3	3	3	2	3
4	6	4	4	4	8	7	8	8	1	3	1	3	3	4	3	5
5	2	2	2	2	9	9	8	9	2	2	2	2	5	5	5	5
6	3	2	4	4	2	4	2	2	3	2	2	2	3	2	0	3
7	4	4	5	6	3	3	2	3	2	2	2	2	4	4	5	4
8	3	0	3	2	7	7	7	7	2	2	2	2	4	4	3	4
Total	51	46	53	50	77	77	70	76	36	36	36	37	38	36	32	38

	Participant 21				Participant 22			
1	18	17	17	18	54	43	12	4
2	7	7	6	8	16	9	5	2
3	4	4	4	4	16	8	6	2
4	3	4	3	3	20	11	5	2
5	4	4	3	2	21	12	4	2
6	7	5	7	6	17	8	6	2
7	2	2	2	2	25	10	7	4
8	2	2	2	2	16	9	7	4
Total	47	45	44	45	185	110	52	22

A nodal number effect was deemed to have occurred on a particular test block when response accuracy was highest on 1-node trial types with a lower response accuracy on 2-node trials, and lowest response accuracy on 3-node trials. Page's Trend test (L) was applied to further test the relatedness, with block number as participant number. Test Blocks on which response accuracy was at 100% for two or more nodal numbers were deemed ineligible because the ceiling effect precluded an analysis. Participants' data were group analyzed according to the mastery criterion as Passed and Failed Participants.

2.2.2.1 Passed Participants

15/21 participants reached the mastery criterion of 90% after repeated exposure to the training and testing cycles, the same stimuli were used for each cycle for each participant. The proportion of the time-out responses in the test of equivalence was similar in all LH conditions, the medians of averages ranged from 1.2% to 2.2%. In contrast, the proportion of the time-out responses in baseline training varied across conditions, with a greater proportion in the Unequal LH condition (median of average: 4.75%) than in the Equal LH condition (median of average: 2%). Table 5 shows the large individual variance (range from 1 to 7) in terms of number of cycles needed for each participant to reach the 90% criterion in the test phase, Participant 7 and 14 reached mastery criterion immediately after initial training phase.

Table 5 Number of cycles needed until participants in each training group reached the mastery criterion of 90% and above accuracy in class consistent responding

Condition	Participant No.	Cycles to criterion
Unequal Reinforcement		
No LH	1	2
No LH	2	6
No LH	3	6
No LH	4	5
LH	5	4
LH	6	6
LH	7	1
LH	8	3
Equal Reinforcement		
No LH	9	2
No LH	10	7
No LH	11	3
No LH	12	3
LH	13	5
LH	14	1
LH	15	3

The data in Table 6 displays the percent correct on unreinforced baseline probes for each participant in test blocks. It showed that only two participants had successfully acquired baseline relations after initial training, while, repeated training and testing required for the rest 13 participants.

Table 6: Percentage correct per test block for unreinforced baseline probes for passed participants

	Test Block Number							
	1	2	3	4	5	6	7	8
Participant	Unequal no LH							
1	62.5	100	100	100	100	87.5	87.5	100
2	87.5	50	87.5	75	100	100	100	100
3	75	75	62.5	75	75	100	100	100
4	75	62.5	87.5	100	100	100	87.5	100
	Unequal LH							
5	87.5	87.5	100	100	100	100	100	87.5
6	37.5	62.5	87.5	75	87.5	100	100	100
7	100	100	100	100	100	100	100	100
8	87.5	87.5	100	100	100	100	100	100
	Equal no LH							
9	75	100	100	100	100	100	100	100
10	87.5	62.5	37.5	87.5	75	100	87.5	100
11	100	100	87.5	100	100	100	100	100
12	75	75	100	100	100	100	100	100
	Equal LH							
13	37.5	87.5	75	87.5	100	100	100	100
14	87.5	100	100	100	100	100	100	100
15	37.5	87.5	100	100	100	100	87.5	100

2.2.2.1.1 Response Speed

The data are expressed as response speed (inverted latency) as this minimizes variance due to long latencies, which are more likely to be due to processes other than those of interest (e.g., due to inattention; Ratcliff, 1993; Whelan 2008). Due to the large within-participant variability of RT, group means are presented. Response speed was analyzed from first block until above 90% accuracy to control the ceiling effect.

Figure 1 displays the mean RS and 95% confidence intervals (CIs) across the four groups. Serial of Page's Trend tests were performed to examine the sequential order effect in the four conditions, a significant nodal effect was found in the Equal reinforcement no LH group only (L=54, p=.05). There appears to be little difference among the RSs for 1-, 2-, and 3-node relations in the Unequal Reinforcement LH, Equal Reinforcement LH, Equal Reinforcement no LH group.

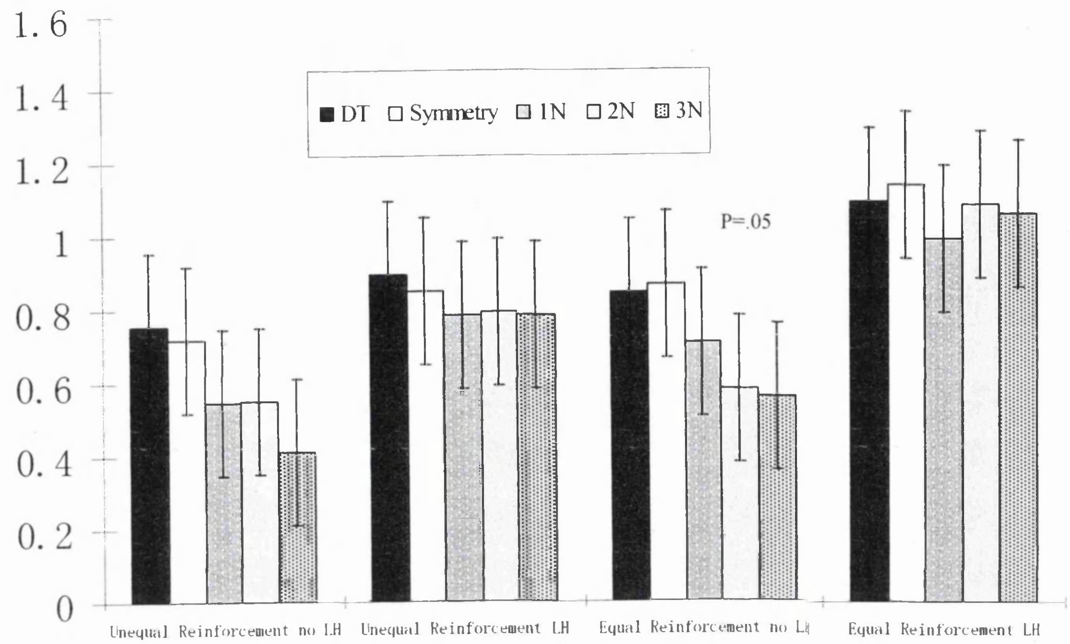


Figure 1: Mean response speed and 95% confidence intervals across all relation types for passed participants. DT: directly trained. 1N: 1 node. 2N: 2 nodes. 3N: 3 nodes

2.2.2.1.2 Response Accuracy

The top panel of Figure 2 displays the results from Unequal Reinforcement no LH condition. Participant 1's response accuracy quickly reached the criterion on the test blocks. On Test Block 1, responses to 2-node trials types were most accurate, followed by responses to 1-node trial types. A nodal number effect emerged immediately after training in Participant 2, and stabilized on Test Blocks 3, 4, and 5 (L=42, p<.05), response accuracy reached the mastery criterion on Test Blocks 6-8 eventually. The responses of Participant 3 showed a reversed order of nodal number effect immediately after training, it seems after self-correction, nodal effect emerged and stabilized in Test Block 4 and 5 (L=28, p=.05) before reached ceiling. The responses of Participant 4 suggest a nodal number effect on two of five eligible blocks (Test Block 4 and 6, L=28, p=.05).

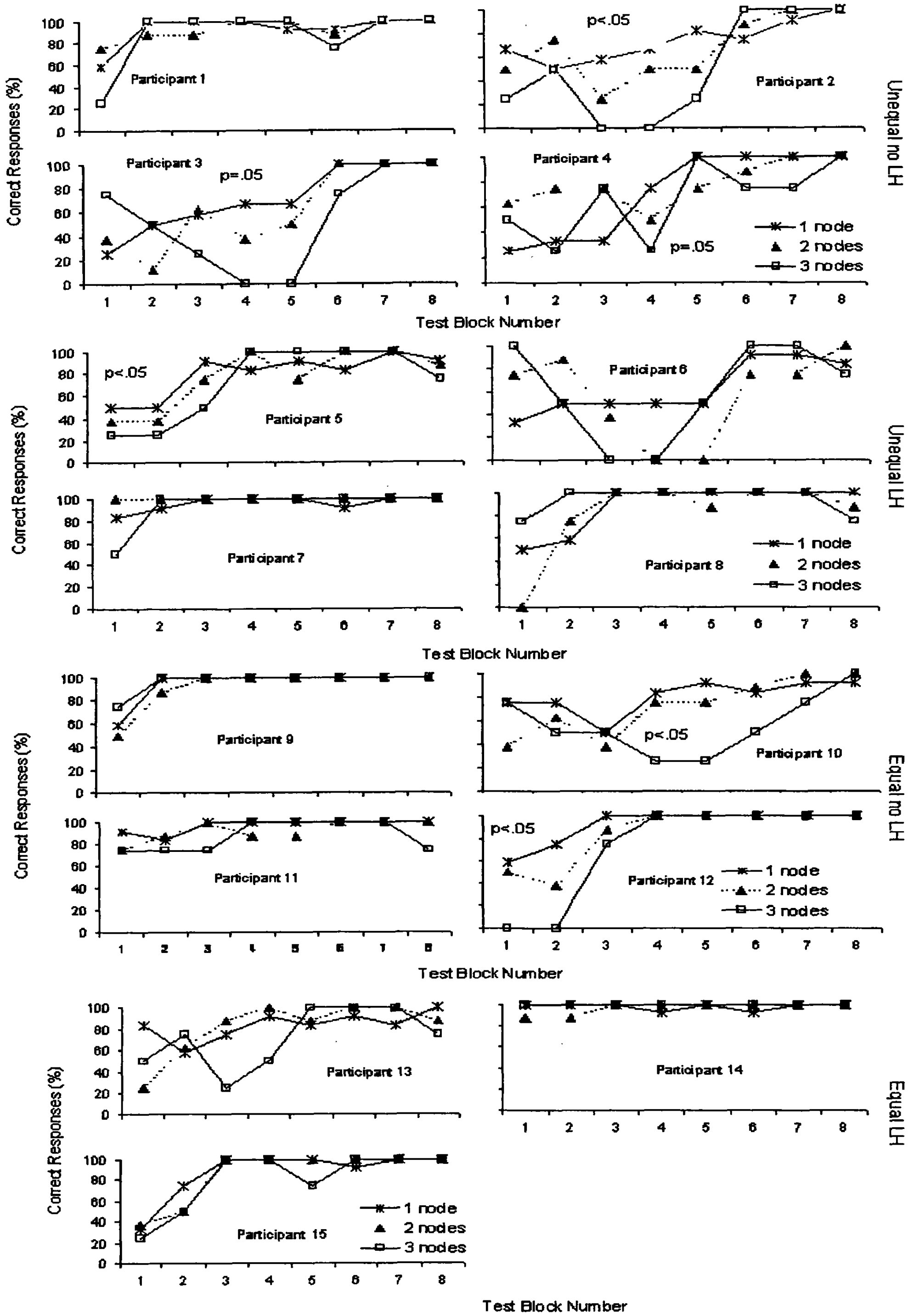


Figure 2: Percent correct for 1-, 2- and 3-nodes across all blocks of testing in all conditions for passed participants.

The second-from-top panel of Figure 2 displays the results from Unequal Reinforcement LH condition. The responses of Participant 5 do suggest an effect of nodal number (L=42, P<.05) across Blocks 1-3. Correct responses on this block were at a very high rate on Blocks 4-8. The responses of Participant 6 suggest a weak effect of nodal number (only Block 3 emerged the nodal effect pattern), interestingly, a reversed order of nodal number effect emerged immediately after training. The response accuracy of Participant 7 was at the mastery within one block, therefore, unable to provide sufficient information for analysis. The response accuracy of Participant 8 was at the mastery within three blocks. A reversed order of nodal effect emerged in the Test Block 2 before reached ceiling in Block 3-7; it reemerged at the last testing block.

The second-from-bottom panel of Figure 2 displays the results from Equal Reinforcement no LH condition. A nodal number effect ($L=42$, $p<.05$) was observed on Blocks 2, 4, and 5 of Participant 10's test blocks. Participant 12 showed a nodal number effect ($L=42$, $p<.05$) on Blocks 1-3, before reaching the mastery criterion on Block 4. The response accuracy of Participant 9 and 11 reached the mastery criterion very quickly, little information was provided.

The bottom panel of Figure 2 displays the results from Equal Reinforcement LH condition. There was no evidence of nodal number effects in this group, with the exception of Participant 13's final test block. The responses of Participant 13 were generally erratic across Blocks 1-5, before stabilizing on Blocks 6-8. The rest participants in this group (pt14 and 15) reached mastery criterion shortly after training.

In summary, 3/4 participants showed nodal effects in certain testing blocks in Unequal no LH condition, interestingly, nodal effects did not disappear under Equal no LH condition, in fact, half of the participants showed nodal effects in certain testing blocks. One participant showed nodal effect in Unequal LH condition, and no nodal effects emerged in Equal LH condition. It would be interesting to examine the performance of failed participants, as a nodality effect seems sometimes to be found even in the absence of equivalence (Fields, et al, 1995; Randall and Remington 1999).

2.2.2.2 Failed Participants

There were 2 participants quitted from Equal Reinforcement LH condition without reaching the mastery criterion of 90% accuracy after repeated exposure to the training and testing cycles. 4 participants quitted in Equal Reinforcement no LH condition, 1 quitted from the Unequal Reinforcement no LH condition, and no participants quitted in Unequal Reinforcement LH condition.

The data in Table 7 displays the percent correct on unreinforced baseline probes for each participant in test blocks. None of them acquired all baseline probes after initial training, 3 participants in no limited hold conditions acquired all baseline probes after repeated training and testing cycles. Greater impairment of performance in Equal Reinforcement LH condition suggested a time-accuracy trade-off.

Table 7: Percentage correct per test block for unreinforced baseline probes for failed participants

	Test Block Number							
	1	2	3	4	5	6	7	8
Participant	Unequal no LH							
22	87.5	100	100	100	100	100	87.5	100
	Equal no LH							
18	87.5	87.5	75	87.5	87.5	87.5	87.5	75
19	75	100	87.5	100	100	87.5	100	100
20	75	75	100	75	100	87.5	100	100
21	75	75	75	50	75	100	100	62.5
	Equal LH							
16	25	50	75	62.5	87.5	62.5	62.5	75
17	37.5	50	75	100	87.5	87.5	62.5	75

2.2.2.2.1 Response Speed

Data treatment was identical to those in passed group (See section 2.2.2.1.1).

Figure 3 displays the mean RS and 95% confidence intervals (CIs) across the three conditions. No significant trend effects were found in all conditions, even though the only participant in Unequal Reinforcement no LH condition showed an obvious reversed linear trend (3N faster than 2N, 2N faster than 1N).

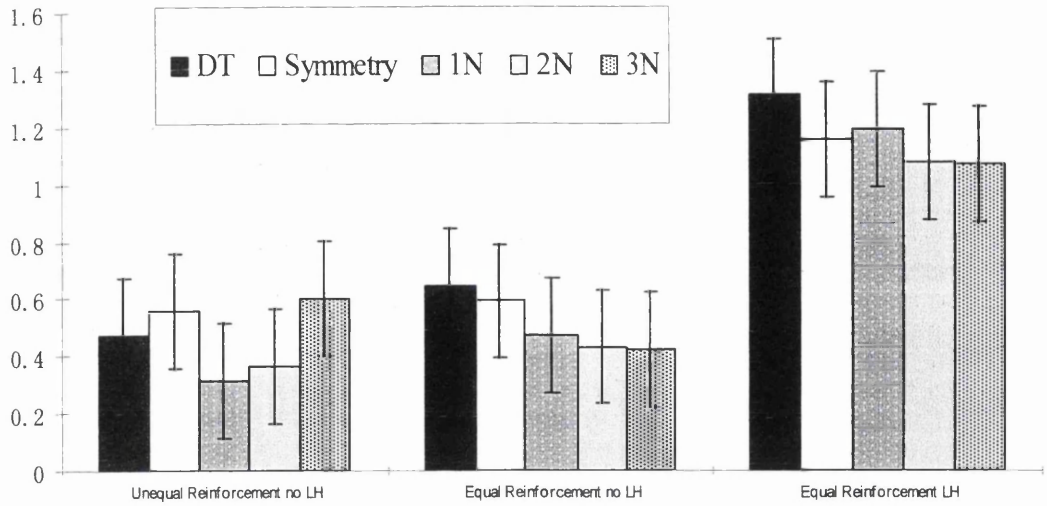


Figure 3: Mean response speed and 95% confidence intervals across all relation types for failed participants. DT: directly trained. 1N: 1 node. 2N: 2 nodes. 3N: 3 nodes

2.2.2.2.2 Response Accuracy

The top panel of Figure 4 displays the results from Equal Reinforcement LH condition. No nodal consistent pattern was observed in this condition. The middle panel of Figure 4 displays the results from Equal Reinforcement no LH condition. Participant 18 performed best on 3N relations, but no difference between 1N and 2N relations. Participant 19 showed an inversed response pattern across test blocks (3N better than 2N, 2N better than 1N). Participant 20 performed best on 3N relations, followed by 1N relations, the least correct responses on 2N relations. Participant 21 performed best on 3N relations, and no consistent patterns observed across test blocks. The lower panel of Figure2.4 displays the only result from Unequal Reinforcement no LH conditions. Participant 22 showed an inversed nodal pattern across test blocks.

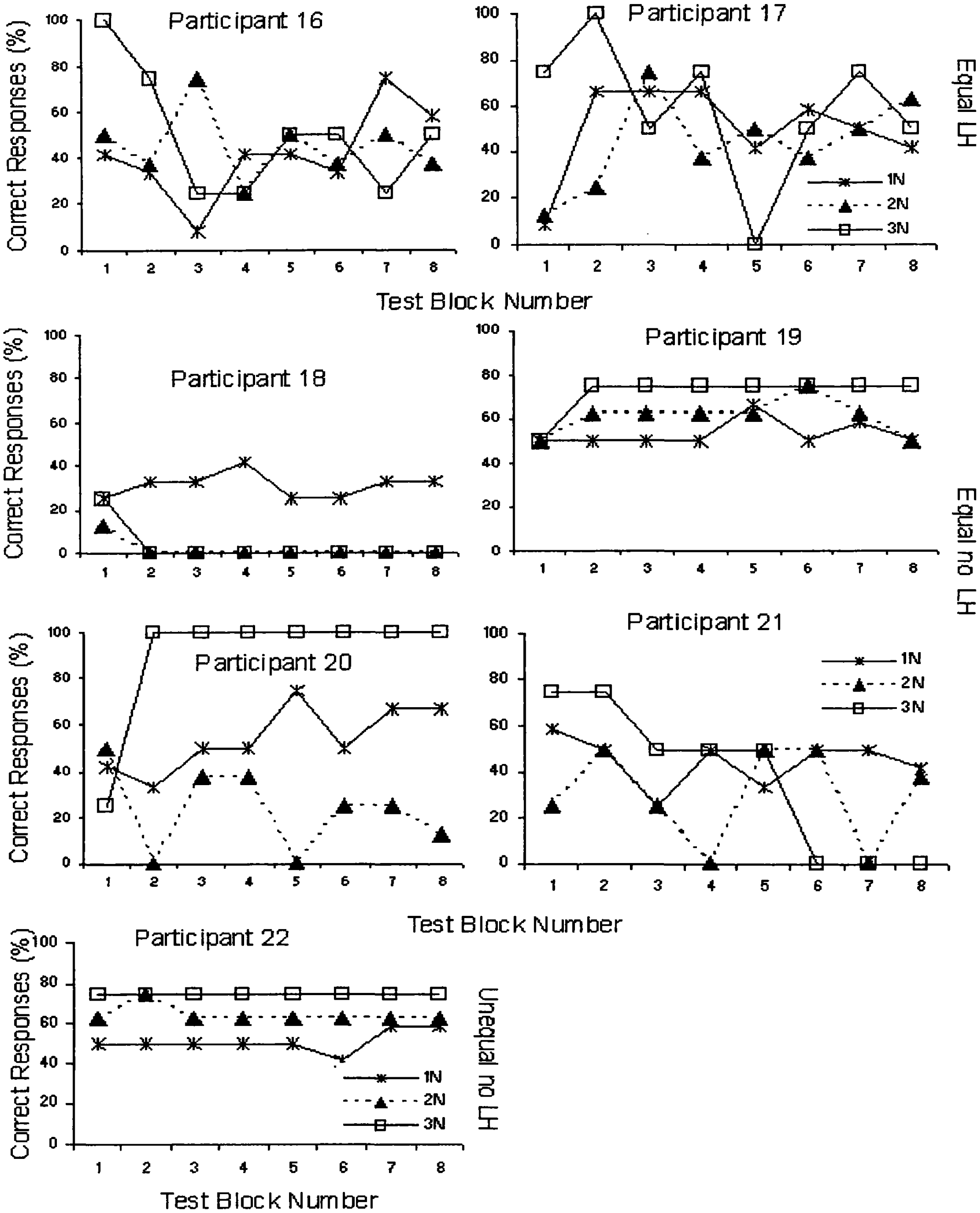


Figure 4: Percent correct for 1-, 2- and 3-nodes across all blocks of testing in all conditions for failed participants.

In summary, No nodal effects were found in analyzing data from RS and RA, even though a reversed nodal pattern often reported in both analyses.

2.2.3 Discussion

In Experiment 1 the number of reinforcers and training structure were manipulated. In the Equal Reinforcement group, the number of delivered reinforcers was approximately equal across all baseline trial types (see Table 3), and all trial types were introduced simultaneously. In contrast, in the Unequal Reinforcement group, the number of delivered reinforcers was very different across trial types (see Table 3), and trial types were introduced serially. The number of participants who eventually reached 90% accuracy for derived relations was higher in the Unequal Reinforcement group than in Equal Reinforcement group. For those participants who passed the test for derived relations, it did not appear to be systematic differences among the groups in terms of the number of cycles needed to reach the mastery criterion for class consistent responding.

The key result in Experiment 1 was that nodal number effect was found in the Equal Reinforcement no LH group for response speed only, thus indicating that serial training structure resulting in unequal reinforcement is *not* a prerequisite for nodal effects to emerge (cf. Imam 2001, Experiment 2). In addition, nodal effects for accuracy were apparent in the *no limited hold* conditions, but not in *limited hold* groups, only with the exception of one participant from the Unequal Reinforcement group. This providing further evidence that a time-accuracy trade off occurred, thus possibly obscuring any nodal number effects (see Dickens, 2005) from observing in *limited hold* groups. In the present study, the start of the test phase was not signaled. The data in Table 5 – present percent correct on unreinforced baseline probes – suggesting that the baseline conditional discriminations were disrupted by the sudden termination of reinforcement.

2.3 Experiment 2

The results of Experiment 1 provided RS evidence to support the prediction that nodality effects are preserved under equal reinforcement, response accuracy indicated that nodal effects did not disappear under equal reinforcement, whereas unequal reinforcement might have a facilitative role in producing nodal effects. Since only 8 participants were recruited under equal reinforcement, Experiment 2 aimed to test whether the observed nodal effects would generalize to a larger sample, therefore, the current experiment employed a similar equivalence training and testing procedure to the Equal Reinforcement protocol in Experiment 1 with an increased sample size. In addition, Experiment 1 confirmed that a speed accuracy trade-off occurred under the LH, therefore a LH was not employed in Experiment 2, as it would likely mask the effect of reinforcement on response performance.

2.3.1 Method

2.3.1.1 Participants

Twenty-three adults participated in the Experiment 2.2 (4 male; 19 female), ranging in age from 18 to 33 years (mean age = 23.45 years, standard deviation = 3.76 years). All participants were students (14 undergraduate, 9 postgraduate) at the University of Wales, Swansea; All participants were recruited through personal contacts by the experimenter (21 Chinese nationals, 1 German national and 1 Greek national) and all were naïve about the purpose of the experiment. The study was approved by the Department of Psychology, University of Wales Swansea, Ethics Committee.

2.3.1.2 Apparatus and Materials

The apparatus, stimuli and setting were the same as Experiment 1.

2.3.1.3 Procedure

The procedure was identical to the Equal Reinforcement no LH in Experiment 1 with the only difference that the experiment was terminated when the participant reached 90% of mastery criterion in the equivalence test. Therefore, the number of cycles needed varied across participants in order to control a ceiling effect.

2.3.2 Results

Twenty-three participants began Experiment 2, twelve of them formed equivalence after repeated exposure to the training and testing cycle.

The actual numbers of reinforcers delivered in Experiment 2 across all participants are presented in Table 8. These data indicate that the procedures employed were successful in manipulating the number of reinforcers across trial types, that is, the number of delivered reinforcers was approximately equal across all trial types for all participants.

Table 8: Number of reinforcers delivered per test block per participant in Experiment 2.

Block	Trial Type															
	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE
	Participant 1				Participant 4				Participant 6				Participant 7			
1	12	12	12	12	19	19	20	20	23	23	23	23	16	16	16	14
2	13	12	12	13	7	8	8	7	24	24	24	24	2	2	2	2
3	2	3	2	2					14	14	14	14				
4									18	16	16	17				
5									29	30	30	30				
Total	27	27	26	27	26	27	28	27	108	107	107	108	18	18	18	16
	Participant 12				Participant 13				Participant 14				Participant 15			
1	16	15	14	14	90	90	90	89	10	9	10	9	8	10	9	9
2	37	37	38	37	44	45	44	45	21	20	20	22	5	5	5	4
3	4	4	4	4	4	4	2	4	2	2	2	2	17	16	16	17
4	2	2	2	2	6	7	6	7					8	6	7	7
Total	59	58	58	57	144	146	142	145	33	31	32	33	38	37	37	37
	Participant 16				Participant 18				Participant 20				Participant 23			
1	27	27	26	26	31	31	30	32	25	24	24	26	18	17	17	18
2	6	8	8	6	5	4	4	4	6	7	7	7	11	11	12	11
3	4	4	4	3	2	3	2	3	13	12	12	12				
4	2	2	2	2	3	3	3	3	32	32	32	32				
5	2	2	2	2	30	29	30	30	6	6	6	6				
6					12	12	10	11								
7					10	12	10	10								
8					10	9	8	9								
Total	41	43	42	39	103	103	97	102	82	81	81	83	29	28	29	29
	Participant 2				Participant 3				Participant 5				Participant 8			
1	4	4	3	3	10	10	10	10	27	28	26	27	48	49	50	49
2	12	12	12	12					6	6	5	6				
3									12	12	12	12				
Total	16	16	15	15	10	10	10	10	45	46	43	45	48	49	50	49
	Participant 9				Participant 10				Participant 11				Participant 17			

1	12	11	12	11	21	21	21	20	7	7	7	7	91	92	91	90
2	17	16	16	17	7	7	8	7	13	12	12	12	11	11	10	10
3									11	12	12	11	2	2	2	2
4									3	4	3	3	10	10	11	11
5									8	9	8	9	2	2	2	2
6													5	6	6	6
7													3	2	2	2
8													11	11	12	12
9													14	14	14	14
10													2	2	2	2
Total	29	27	28	28	28	28	29	27	42	44	42	42	151	152	152	151
	Participant19				Participant 21				Participant22							
1	6	5	5	5	4	4	4	3	14	14	15	16				
2	7	6	6	6	92	93	92	92	18	18	16	18				
3	2	2	2	2	9	9	9	10								
4	8	8	9	9												
5	10	11	10	11												
6	12	13	12	12												
Total	45	45	44	45	105	106	105	105	32	32	31	34				

2.3.2.1 Passed Participants:

Table 9 shows the Cycle number on which each participant reached the mastery criterion in the equivalence test. All of them completed at least 2 cycles of training and testing.

Table 9: Number of cycles needed until participants in each training group reached the mastery criterion of 90% and above accuracy in class consistent responding in Experiment 2

Participant No.	Cycles to Criterion
Equal Reinforcement no LH	
1	3
4	2
6	5
7	2
12	4
13	4
14	3
15	4
16	5
18	8
20	5
23	2

The data in Table 10 displays the percent correct on unreinforced baseline probes for each participant in test blocks. 4 participants acquired all baseline probes immediately after initial training. Repeated training and testing is required for the rest 8 participants.

Table 10: Percentage correct per test block for unreinforced baseline probes for passed participants

	Test Block Number							
	1	2	3	4	5	6	7	8
Participant	Equal no LH							
1	62.5	75	100					
4	100	100						
6	50	37.5	37.5	62.5	100			
7	100	100						
12	62.5	75	87.5	100				
13	37.5	50	100	100				
14	87.5	100	100					
15	100	62.5	100	100				
16	87.5	87.5	100	87.5	100			
18	100	100	75	87.5	100	87.5	62.5	100
20	87.5	62.5	62.5	100	100			
23	37.5	100						

2.3.2.1.1 Response Speed

Figure 5 displays the mean RS and 95% confidence intervals (CIs) across the three conditions. No significant trend effects were found in all trial types.

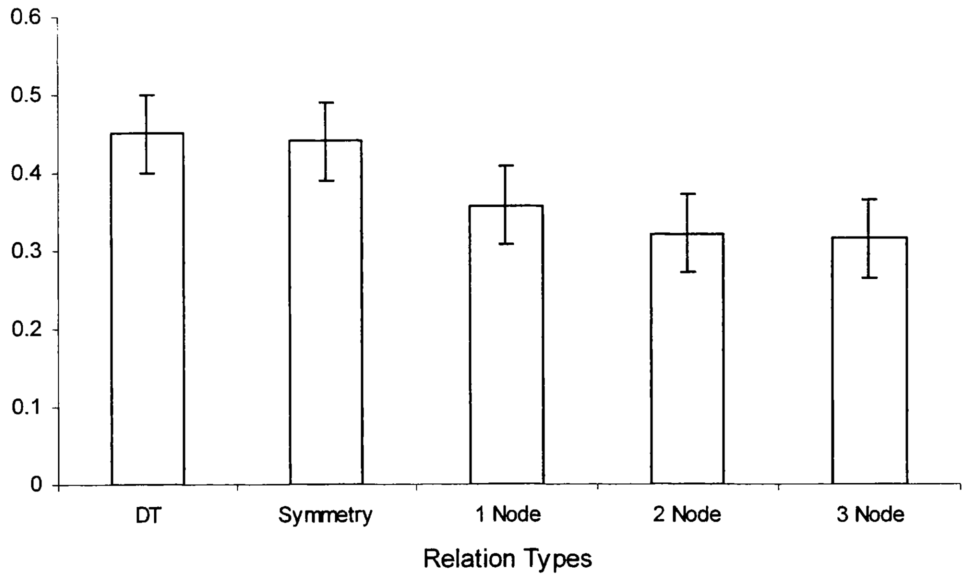


Figure 5: Mean response speed and 95% confidence intervals across all relation types for passed participants. DT: directly trained. 1N: 1 node. 2N: 2 nodes. 3N: 3 nodes

2.3.2.1.2 Response Accuracy

Percent of response accuracy for each participant in terms of the nodality hypothesis across the number of test blocks is shown in figure 6. The top panel of the Figure 6 displays the results from participant 1, 4, 6 and 7. Participant 1's response accuracy indicated a nodality trend cross block 1 and 2 in terms of responding to 1 and 2-node trial types, however, in block 2, responses to 3-node trial types were most accurate. Participant 4's response accuracy quickly reached the criterion on the test blocks, and no difference in responses to derived trial types was observed. Clear nodality effects were demonstrated on Test Blocks 2 and 4 for the responses of Participant 6, responses to 3-node trial types were least accurate across all test blocks. Participant 7's response accuracy quickly reached the criterion on the test blocks, with a ceiling effect emerging on 1and 2-node trial types in test blocks.

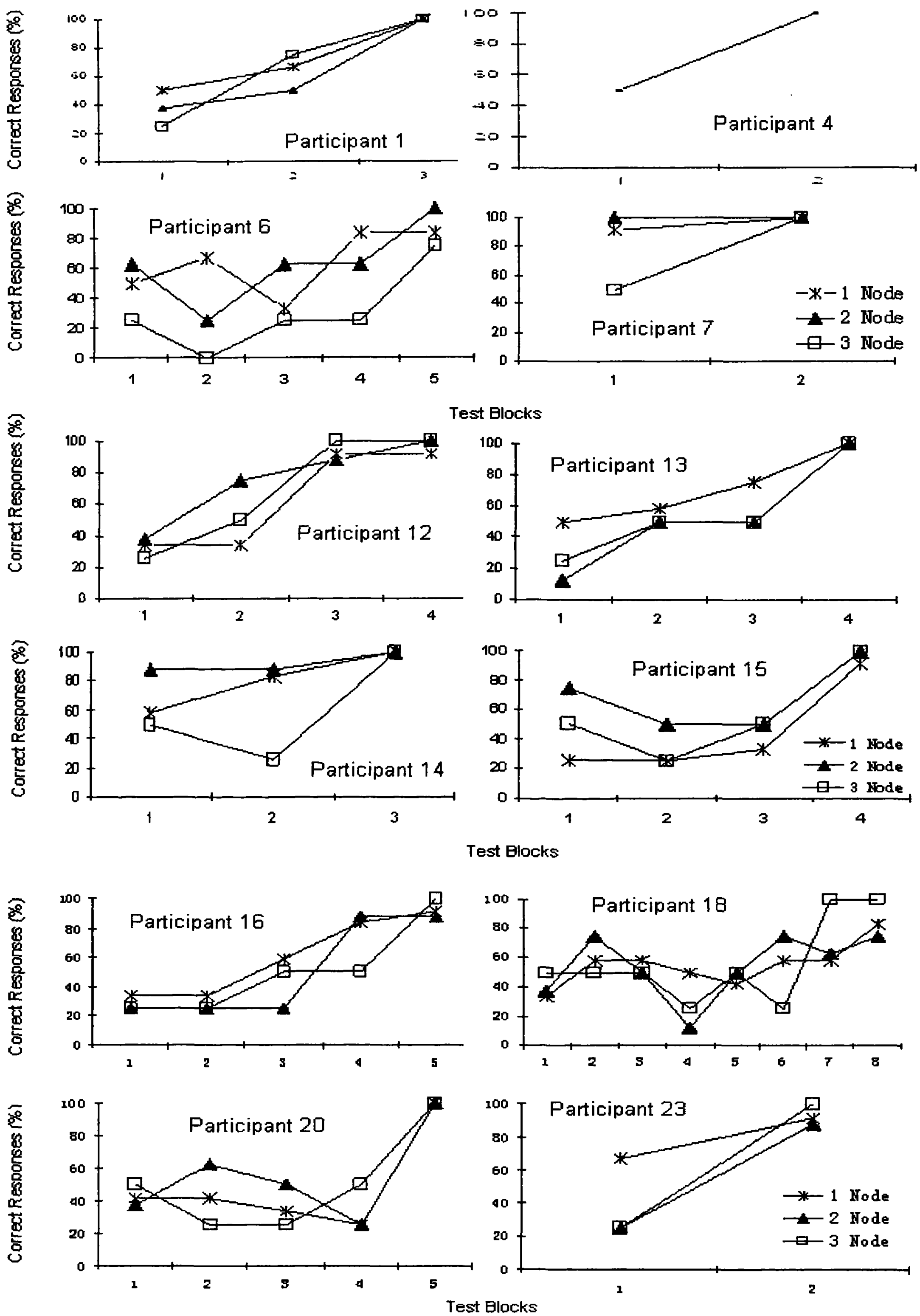


Figure 6: Percent correct for 1-, 2- and 3-nodes across all blocks of testing in all conditions for passed participants

The middle panel of Figure 6 displays the results from participant 12, 13, 14 and 15. Participant 12's response accuracy reached ceiling at Test Block 3 and 4, no nodality effect was observed. A nodality effect was apparent in the first test block for the responses of Participant 13, responses to 2 and 3-node trial types yield equal accuracy in remaining blocks before reaching the mastery criterion. Responses to 1-node trial types were most accurate across test blocks. Although no apparent nodality effect was demonstrated across test blocks on responses of Participant 14, responses to 3-node trial types were least accurate in Test Block 1 and 2, before reaching the mastery criterion. Participant 15's response accuracy was contrary to the nodality hypothesis, with responses to 1-node trial types least of accurate across test blocks.

The bottom panel of Figure 6 displays the results from participants 16, 18, 20 and 23. No nodality effect was observed from responses of Participant 16. The responses of Participant 18 were erratic across test blocks, and difficult to interpret. Participant 20's response accuracy was not consistent with the nodality hypothesis before reaching the mastery criterion on Block 5. Participant 23's response accuracy reached the mastery criterion after exploring to 2 test blocks, with responses to 1-node trial types the most accurate in the first test block.

2.3.2.2 Failed Participants

11/23 participants quitted the experiment without reaching the mastery criterion of 90% correct in equivalence test. Participant 8 quitted without completing a single test block, therefore, no teat data obtained for this participant. The data in Table 11 displays the percent correct on unreinforced baseline probes for each participant in test blocks. 4 participants acquired all baseline probes after repeated training and testing cycles, in which Participant 3 acquired all baseline probes immediately after initial training.

Table 11: Percentage correct per test block for unreinforced baseline probes for failed participants

	Test Block Number									
	1	2	3	4	5	6	7	8	9	10
Participant	Equal no LH									
2	62.5	87.5								
3	100									
5	75	100	62.5							
9	50	37.5								
10	75	100								
11	62.5	62.5	100	100	100					
17	62.5	87.5	62.5	87.5	87.5	62.5	87.5	87.5	87.5	87.5
19	75	100	87.5	100	100	100				
21	62.5	75	62.5							
22	75									

2.3.2.2.1 Response Speed

Figure 7 displays the mean RS and 95% confidence intervals (CIs) across the three conditions. No significant trend effects were found across all trial types.

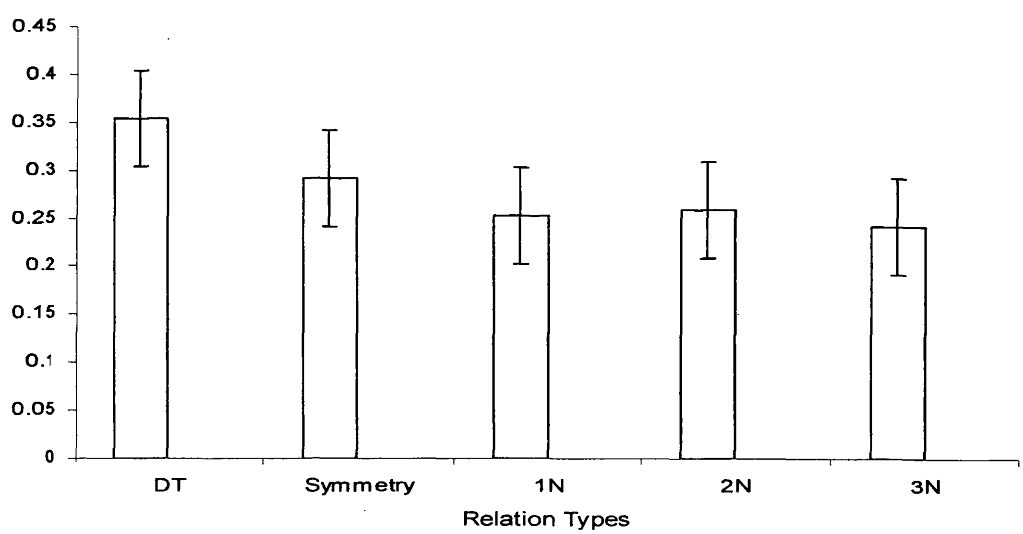


Figure 7: Mean response speed and 95% confidence intervals across all relation types for failed participants. DT: directly trained. 1N: 1 node. 2N: 2 nodes. 3N: 3 nodes

2.3.2.2.2 Response Accuracy

Percent of response accuracy for each participant in terms of the nodality hypothesis across the number of test blocks is shown in figure 8. No nodality consistent patterns found across participants, although Participant 3 and 21 demonstrated an inversed pattern of nodal effect (3N better than 2N, 2N better than 1N).

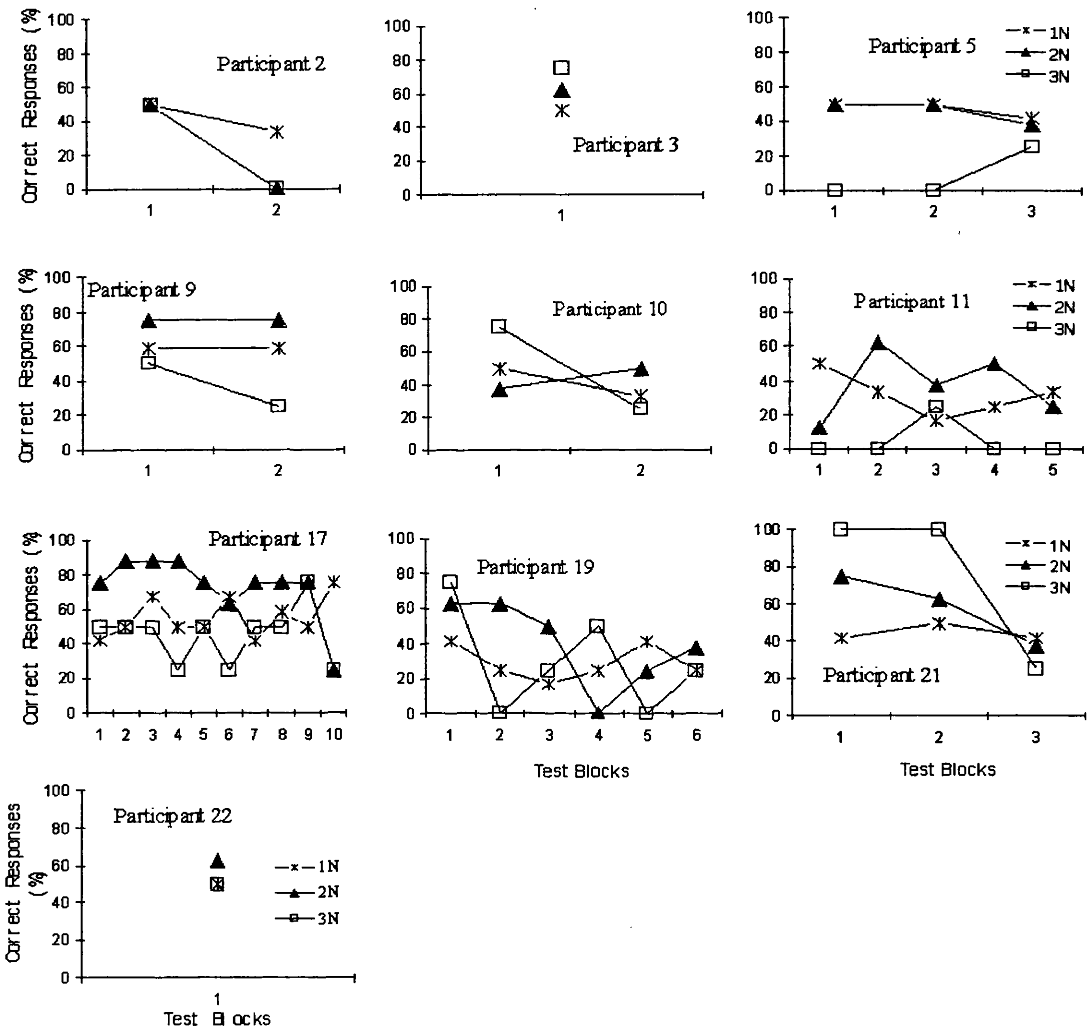


Figure 8: Percent correct for 1-, 2- and 3-nodes across all blocks of testing.

2.3.3 Discussion

Interestingly, no sound nodal effects were found in the analysis of both RS and RA in Experiment 2, which contradict the findings in Experiment 1. This discrepancy might be due to the disrupted baseline probes in tests of equivalence resulting in poor equivalence performance. According to Fields and his colleagues (2007, 2008), the format of MTS test used to establish equivalence performance maximizes the between class discrimination rather than the node based within class discrimination. Therefore, test format that maximize within class discrimination would account for the emergence of nodal effects in some studies but not others. Experiment 3 aims to address this issue.

2.4 Experiment 3

Although the results of Experiment 1 provided evidence for the prediction that nodal number effects are preserved under equal reinforcement and a simultaneous training protocol in response speed, opposing findings were produced in Experiment 2. The different results obtained from Experiment 1 and 2 may be due to disrupted baseline probes in testing and the format of MTS test used to establish equivalence performance. A more sophisticated approach was employed by Fields et al., 1993, 1995, to investigate nodality effects involving a *transfer paradigm.* Specifically, their transfer paradigm involved training a function to a particular stimulus or stimuli in a class, while observing the degree to transfer to other class members. According to Fields et al. (1993), "...if the degree of transfer was a systematic function of a variable such as nodal distance...that variable would account for the relatedness of the stimuli in the class" (p. 86).

Experiment 3 employed a similar equivalence training and testing protocol to Experiment 1, with the addition of a response transfer test after the test phase and the expansion of class size to six members, criterion of proceeding to the test phase was increased to 10 consecutively correct trials in conditional discrimination training to stabilize performance. Experiment 1 confirmed that a time accuracy trade-off occurred under the *Limited Hold conditions* (LH): therefore a LH was not employed in Experiment 3, as it would likely mask the nodal number effect on response accuracy or transfer of function test performances. In the function-training phase, differential responses were trained to the C and D stimuli in each class using corrective feedback. Next, the A, B, E, and F stimuli were presented in the absence of corrective feedback and the number of responses to each was observed. If the prediction that equal reinforcement will eliminate a nodal effect is correct, then responses to A, B, E and F should be distributed equally following Equal Reinforcement training. In contrast, if the nodal account is correct, then the A and B stimuli should evoke the response trained to the C stimulus, and the E and F stimuli should evoke the response trained to the D stimulus, despite the differential reinforcement. In addition, if unequal reinforcement is indeed a confounding variable then, following unequal reinforcement training, responses trained to the C stimulus should transfer to the A and B stimuli more readily than responses to the D stimulus transfer to the E and F stimuli. That is, B-C and A-B relations should be more strongly established, whereas D-E and E-F relations should be weak.

2.4.1 Method

2.4.1.1 Participants

Eight participants began Experiment 3 (5 male; 3 female), ranging in age from 21 to 29 years (mean age = 24 years, standard deviation = 2.4 years). All participants were students (1 undergraduate, 7 postgraduate) at Swansea University. Participants were recruited through personal contacts by the first author (2 British nationals and 6 Chinese nationals). All were naïve about the purpose of the experiment. The study was approved by the Department of Psychology, Swansea University, Ethics Committee.

2.4.1.2 Apparatus and Materials

The apparatus, setting were identical to those employed in Experiment 1. Two six-member equivalence classes were trained to accommodate the *transfer paradigm*. All trial types were presented in Table 12.

Table 12: Trial types per relation type that were presented in Experiment 3.

Relation Type	Trial Type							
Directly Trained								
	A1B1	B1C1	C1D1	D1E1	E1F1			
	A2B2	B2C2	C2D2	D2E2	E2F2			
Symmetry								
	B1A1	C1B1	D1C1	E1D1	F1E1			
	B2A2	C2B2	D2C2	E2D2	F2E2			
1 Node								
	A1C1	B1D1	C1E1	D1F1	F1D1	C1A1	D1B1	E1C1
	A2C2	B2D2	C2E2	D2F2	F2D2	C2A2	D2B2	E2C2
2 Node								
	A1D1	B1E1	C1F1	D1A1	B1E1	F1C1		
	A2D2	B2E2	C2F2	D2A2	B2E2	F2C2		
3 Node								
	A1E1	B1F1	E1A1	F1B1				
	A2E2	B2F2	E2A2	F2B2				
4 Node								
	A1F1	F1A1						
	A2F2	F2A2						

2.4.1.3 Procedure

Consent form is identical to the one used in Experiment 1. All participants were exposed to the experimental procedure individually across a number of sessions, each lasting between 20-35 mins and scheduled in 2 days.

Each participant was randomly assigned to one of two conditions that differed based on level of Reinforcement (Equal vs. Unequal). The procedure was broadly similar to Experiment 1, with the following exceptions. In the Unequal Reinforcement condition two 6-member equivalence classes were established by training AB, BC, CD, DE and EF relations in a serial manner. The criterion to proceed to the next training phase, or to the test phase, was 10 consecutively correct responses. In the Equal Reinforcement condition, the two equivalence classes were established by training AB, BC, CD, DE and EF relations on a simultaneous basis, so that each relation was presented the same number of times. The criterion to proceed to the test phase was 10 consecutively correct responses.

Once the criterion for the training session had been met, the test phase commenced and the corrective feedback terminated. All baseline conditional relations, tests for mutual entailment, 1-, 2-, 3- and 4-node transitivity and equivalence trials were presented in a single block. The mastery criterion for testing was at least 90% class consistent selection across the block of 60 test trials. The criterion for progressing to the function training phase was originally defined as two consecutively correct test blocks. However, for Subjects 17, 18, 20, 22, 23 the criterion was accidentally set at three consecutively correct test blocks. Upon reaching the criterion across either two or three test blocks, participants were immediately exposed to the function training.

The function training and response transfer testing were conducted entirely by means of the computer and began with the presentation of the following instructions on the computer monitor (adapted from Fields et al., 1995).

> "In this phase, each spacebar press will produce a brick on the screen. Look at the word at the top of the screen. Your task is to learn how many bricks you should build, either 3, 5, 7 or 9 depending on what word is displayed at the top. Press the "Finish" button when you want to complete a trial. You may start a trial again, if you wish, by pressing the "Start Again" button. Sometimes you will receive feedback and sometimes you will not. Please try your best on all tasks."

The instructions cleared when a button with the caption "Press to start", which was underneath the statement, was pressed. In the training phase, two members from each equivalence class functioned as discriminative stimuli (S^Ds). The stimuli were identical to those in the equivalence training and testing phases. Each S^D was presented on the top-centre of the screen of the monitor against a white background. Pressing the spacebar produced a picture of a brick, which appeared at the center-bottom of the screen. Each brick was a dark red rectangle (1 cm in width and 5 cm in length). Clicking the red "Start Again" button at the left bottom of the screen made all bricks on the screen disappear, and set the response counter to zero. Clicking the green "Finish" button at the right bottom of the screen produced corrective feedback ("Correct" or "Wrong", identical to the equivalence training phase) followed by a 1.5-

s ITI (a blank screen) during the training stage, or only the 1.5-s ITI during the test phase(see Illustration 3). The objective was to create 3, 5, 7, or 9 bricks depending on the stimulus displayed at the top of the screen. If a participant made more than 12 responses, the bricks disappeared and began forming again on the bottom of the screen.

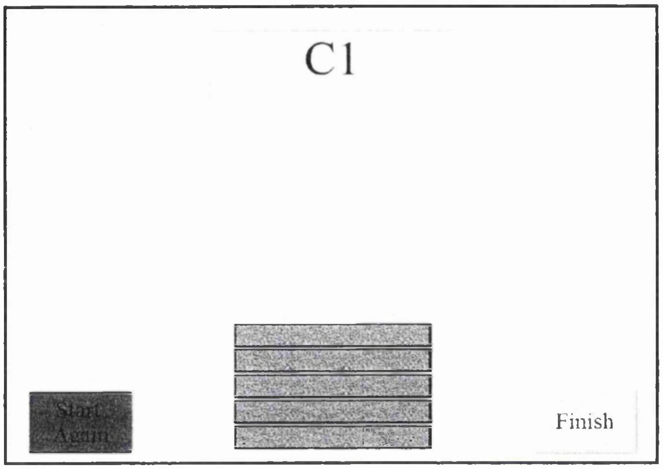


Illustration 3: A screenshot of the function training in Experiment 3.

The following responses were reinforced during the discrimination training: producing three bricks in the presence of the C1 stimulus, five bricks in the presence of the C2 stimulus, seven bricks in the presence of the D1 stimulus, and nine bricks in the presence of the D2 stimulus. Feedback was presented on all trials until the training criterion – eight consecutively accurate responses – was reached, whereupon feedback was stopped without warning. A 72-trial test block was then presented in which each of the six stimuli from each equivalence class was presented six times.

Following the first function transfer phase, each participant was re-exposed to equivalence training to criterion. The equivalence test phase was then presented, and upon passing this test participants were re-exposed to function transfer training and testing. The experiment was concluded following this second function transfer test. Each participant was then thanked for participating, and was debriefed.

2.4.2 Results

Eight participants began Experiment 3; four passed the Unequal Reinforcement condition, four passed the Equal Reinforcement condition. Table 13 shows the number of cycles needed for each participant to reach the mastery criterion in the test phase.

Table 13: Number of cycles needed until participants in each training group reached the mastery criterion of 90% and above accuracy in class consistent responding in Experiment 3.

Participant No.	Cycles to criterion
Unequal Reinforcement	
16	1
17	4*
18	2
19	1*
Equal Reinforcement	
20	4
21	3
22	2
23	2

(*) first two cycles were missing

The actual numbers of reinforcers delivered in Experiment 3 across all participants are presented in Table 14. These data indicate that the procedures employed were successful in manipulating the number of reinforcers across trial types in both conditions. In the Unequal Reinforcement condition, the most number of reinforcers were delivered during AB trials, the next highest amount during BC trials, then CD trials, then DE trials, and the lowest number of reinforcers during EF trials. In contrast, the number of delivered reinforcers was approximately equal across all trial types in the Equal Reinforcement condition, with the exception of Participant 21. In the case of Participant 21, the number of reinforcers delivered was approximately equal for AB, BC and CD trial types, whereas the number of delivered reinforcers was fewer for DE and EF trial types.

Table 14: Number of reinforcers delivered per training block per participant in Experiment 3.

Block	Trial Type																			
	AB	BC	CD	DE	EF	AB	BC	CD	DE	EF	AB	BC	CD	DE	EF	AB	BC	CD	DE	EF
	Participant 16					Participant 17					Participant 18					Participant 19				
1	31	18	12	6	3	22	14	8	4	2	34	19	11	7	4	65	28	9	5	1
2	29	14	8	5	3	29	13	8	6	2	25	19	10	7	1	22	13	7	5	3
3						22	14	8	4	2	21	15	8	4	2					
4						20	12	10	6	2	25	17	9	5	2					
5						24	12	8	5	2										
6						21	13	10	4	2										
Total	60	32	20	11	6	138	78	52	29	12	105	70	38	23	9	87	41	16	10	4
	Participant 20					Participant 21					Participant 22					Participant 23				
1	18	14	19	13	14	49	44	50	15	19	37	40	32	23	23	6	6	7	5	10
2	2	4	4	5	4	8	8	8	4	4	12	8	13	4	9	8	11	7	11	9
3	2	2	2	2	2	15	21	21	19	15	2	2	2	2	2	2	2	2	2	2
4	2	2	2	2	2	1	3	2	2	2	2	2	2	2	2	2	2	2	2	2
5	2	2	2	2	2															
6	2	2	2	2	2															
Total	28	26	31	26	26	73	76	81	40	40	53	52	49	31	36	18	21	18	20	23

A nodal number effect was deemed to have occurred on a particular test block when response accuracy was highest on 1-node trial types, lower on 2-node trials, lower again on 3-node trials, lowest on 4-node trials.

Table 15 displays the percent correct on unreinforced baseline probes for each participant during the test blocks. Participants from the Equal reinforcement group demonstrated much more disrupted baseline conditional discriminations than the Unequal reinforcement group immediately after the first training session.

Table 15: Percentage correct per test block for unreinforced baseline probes for Experiment 3.

	Test Block Number								
	1	2	3	4	5	6	7	8	9
Participant	Unequal Reinforcement								
16	100	100	100						
17	xxx	xxx	100	100	100	100	100	100	100
18	90	100	100	100	80				
19	xxx	xxx	100	100	100				
	Equal Reinforcement								
20	60	100	90	100	100	100	100		
21	60	100	90	100	100				
22	60	100	100	100	100	100	100		
23	40	100	100	100	100	100	100		

(xxx) missing test block.

2.4.2.1 Response Speed

The same data treatment was employed as in Experiment 1 and 2. Figure 9 depicts mean RSs and 95% CIs across all relation types for all participants in each condition for Experiment 3. No significant trend effect was found in either condition.

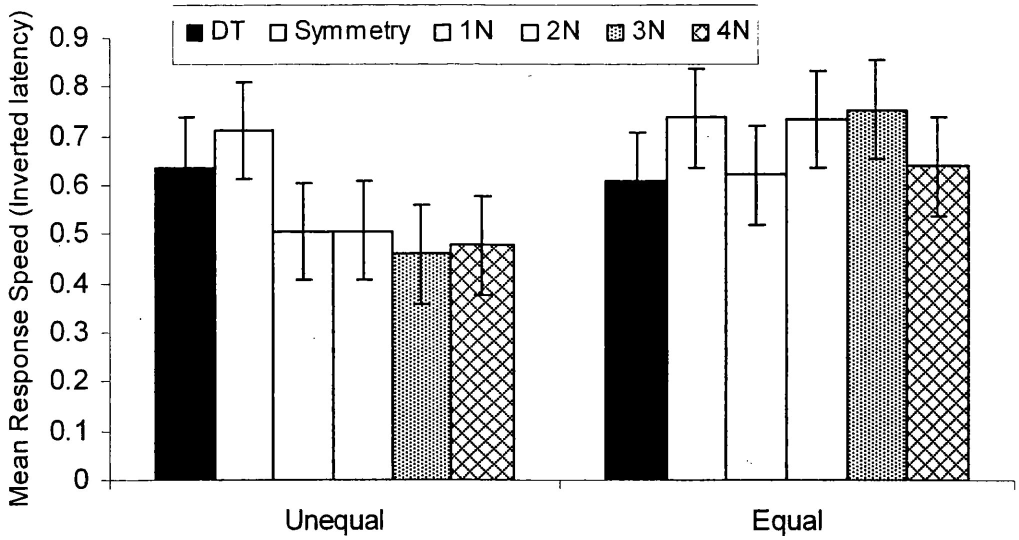


Figure 9: Mean response speed (Inverted latency) and 95% confidence intervals across all relation types for all participants in each condition in Experiment 3. DT: directly trained. 1N: 1 node. 2N: 2 nodes. 3N: 3 nodes. 4N: 4 nodes.

2.4.2.2 Response Accuracy

The upper panel of Figure 10 shows the responses of participants in the Unequal Reinforcement condition in the test phase. The data from the first two cycles of Participants 17 and 19 were lost due to a computer hard-disk failure, and thus accuracy scores on Test Blocks 1 and 2 could not be graphed for these participants. Half of the participants (pt16, and 19) reached ceiling performance shortly after training, rendering it difficult to draw conclusions based on their accuracy data. For Participant 17 from the Unequal Reinforcement group, a significant nodal number effect (defined as highest percent correct on AB trials, then on BC, CD, DE and finally on EF trials) was found on Blocks 3 and 4, according to Page's Trend test ($L=60$, $p<.05$). Participant 18 reached ceiling at Block 2, and showed no sign of node effect.

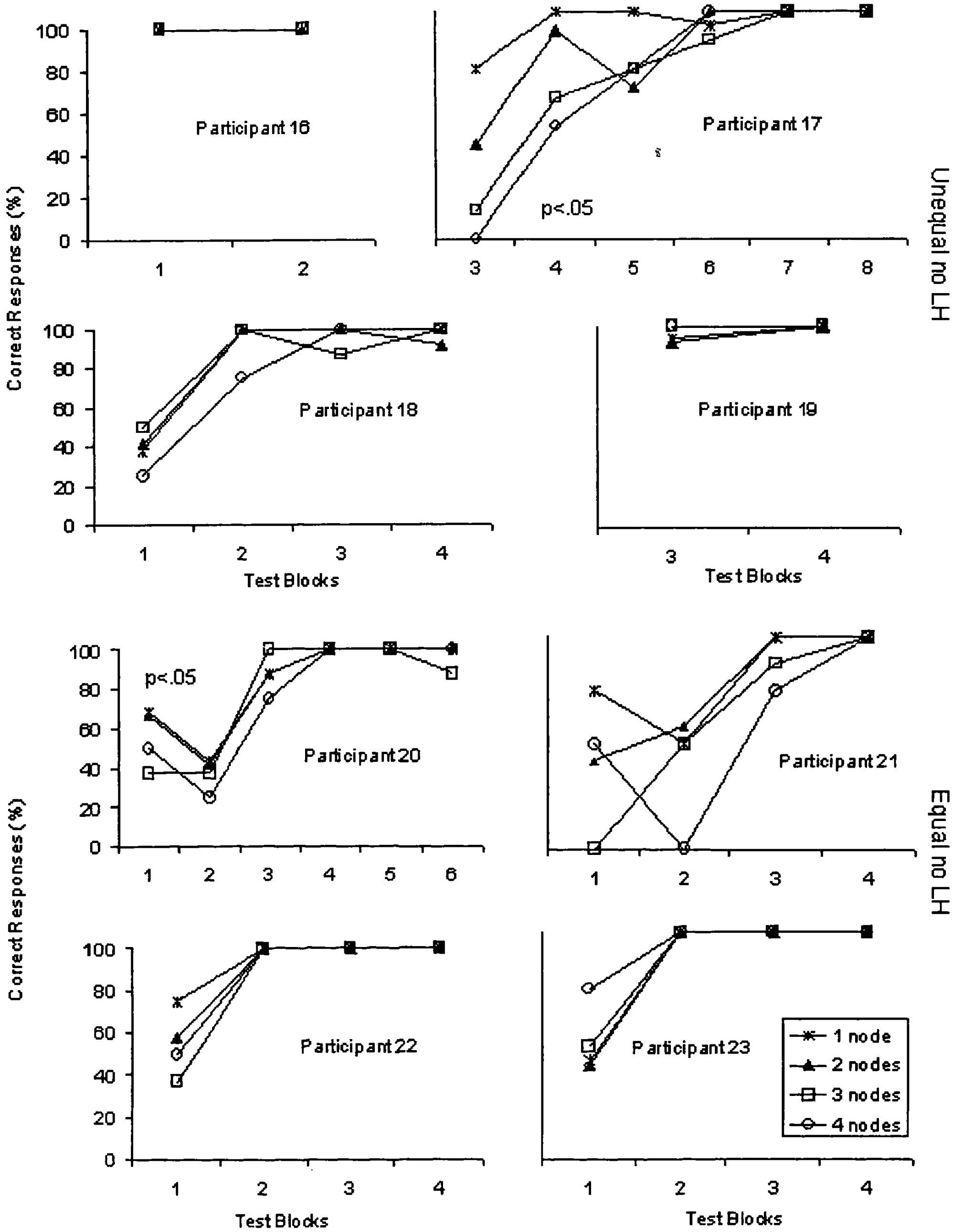


Figure 10: Percent correct for 1-, 2- and 3-nodes across blocks of testing in all conditions in Experiment 3.

The lower panel of Figure 10 shows the responses of participants in the Equal Reinforcement condition in the test phase. Participant 20 demonstrated a significant nodal effect on Block 1 and 2 (L=59, p<.05). Participant 21's response pattern only considered coherent with nodal number prediction if 4-node trials are excluded in Test Block 1 (1 node > 2 node > 3 node), and partially consistent in Test Block 3 (1 node = 2 node > 3 node > 4 node). Participant 22's response pattern partially consistent with nodal effect in the first testing block (1 node > 2 node > 4 node > 3 node), then reached ceiling from Block 2. Participant 23's response showed an interesting reversed ordering of nodal effect (4 node > 3 node > 2 node = 1 node) before reached ceiling at Block 2.

2.4.2.3 Transfer Test

Response transfer was assessed by measuring the relative frequency with which the responses trained to the C and D stimuli were evoked by the other experimental stimuli. For example, the C1 stimulus was an S^D for 3 spacebar presses; if 3 spacebar presses were evoked during all six presentations of the A1 stimulus then the relative frequency was 100%. In some cases, the sum of responses trained to the A, B, E or F stimuli in a class did not equal 100%. This occurred because some responses other than those trained to the C and D stimuli were evoked (e.g. across-class errors). If the response trained to a C or a D stimulus was evoked by other members of the same class, then the response function was deemed to have transferred in accordance with an equivalence relation.

More importantly, if the response trained to C was evoked by the A and B stimuli, and not by the E and F stimuli, within the same equivalence class, then those responses were deemed to be also under the control of nodal number (similarly, if the response trained to D was evoked by the E and F stimuli and not by the A and B stimuli). Thus, a series of exact tests were performed on C and D stimuli in both the serial and simultaneous conditions to further examine the differences between the proportion of responses for the C and D stimuli. If the proportion of responses that transferred to the A and B stimuli was not the same as those that transferred to the E and F stimuli, the null hypothesis would be rejected. In general, presentation of C and D stimuli in each class almost always occasioned the trained responses, thus discriminative control by the C and D stimuli was maintained in the absence of explicit reinforcement.

The proportion of transferred responses was analyzed based on condition. In the Unequal reinforcement condition, participants' proportion of responses that transferred to the A1 and B1 stimuli was significantly greater than those on the E1 and F1 stimuli, when trials were controlled by the C1 stimulus ($p=.0315$, one-tailed); was significantly greater on E1 and F1 stimuli when trials were controlled by D1 stimulus ($p=.008$, one-tailed). No significances were found for the C2 and D2 stimuli. In the Equal Reinforcement condition, participants' proportion of transferred responses was significantly greater on A2 and B2 stimuli when trials were controlled by the C2 stimulus (p=.0155, one-tailed); and was significantly greater on E2 and F2 stimuli when trials were controlled by D2 stimulus (p=.0155, one-tailed). No significant differences were found on the C1 and D1 stimuli. Exact tests suggested that responses transfer occurred amongst one equivalence class members in both condition.

The analysis of individual participant's response patterns in the Unequal Reinforcement condition are shown in Figure 11. The responses of Participant 16 were not controlled by nodal number. For example, the E1 stimulus did not evoke the response trained to the D1 stimulus, but rather evoked the response trained to the C1 stimulus. In contrast, the results of Participant 17 indicate that untrained responses were evoked according to nodal number because the E and F stimuli evoked the response trained to the D stimulus equally often as the A and B stimuli evoked the response trained to the C stimulus. The response pattern of Participant 18 was similar to those of Participant 17, although the relative frequency of responding is slightly below 100% in the presence of the D1 stimulus in Session 1. Participant 19's data show that A1 and B1 did not evoke any trained responses (i.e., responses other than those trained to C1or D1 were evoked). E1 and F1 and F1 in Session two evoked the D1 response. The C2 response was evoked in the presence of B2 and F2 in both Session 1 and 2. No D2 responses were evoked by any of the Class 2 stimuli.

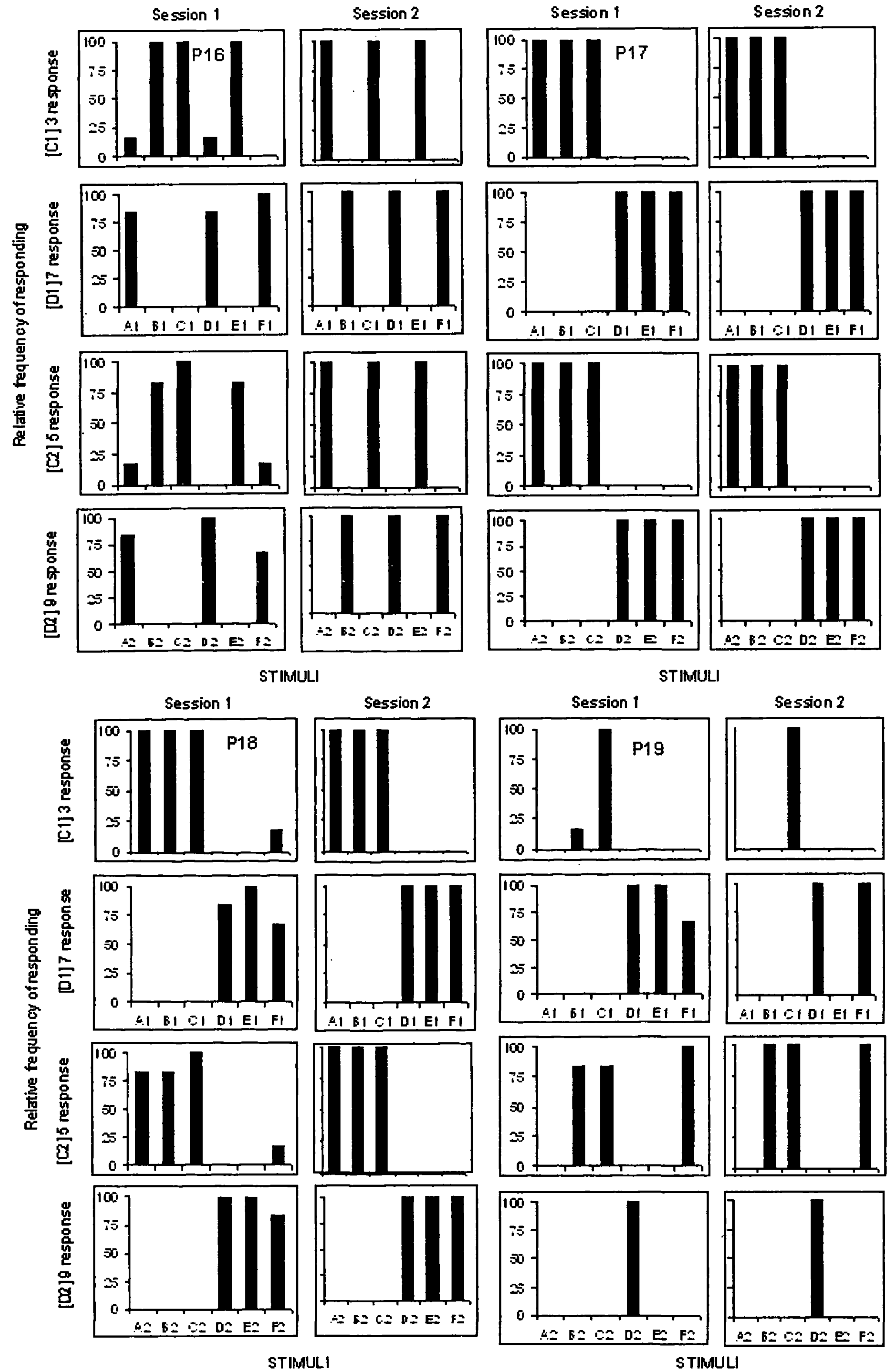


Figure 11: Relative frequency with which responses trained to C and D were evoked by stimuli in both classes across two sessions in the Unequal Reinforcement group in Experiment 3

In the Equal Reinforcement condition (see Figure 12), the responses of Participants 20 and 21 do not appear to be under the control of nodal number, with responses distributed among members of the equivalence class. The response patterns of Participants 22 for Class 1 indicate that responding was under the control of nodal number. Responses to untrained Class 2 stimuli show a slightly weaker nodal effect, with not all untrained stimuli evoked the trained response. The results of Participant 23 are similar, with the F2 stimulus evoking the C2, and not the D2, response.

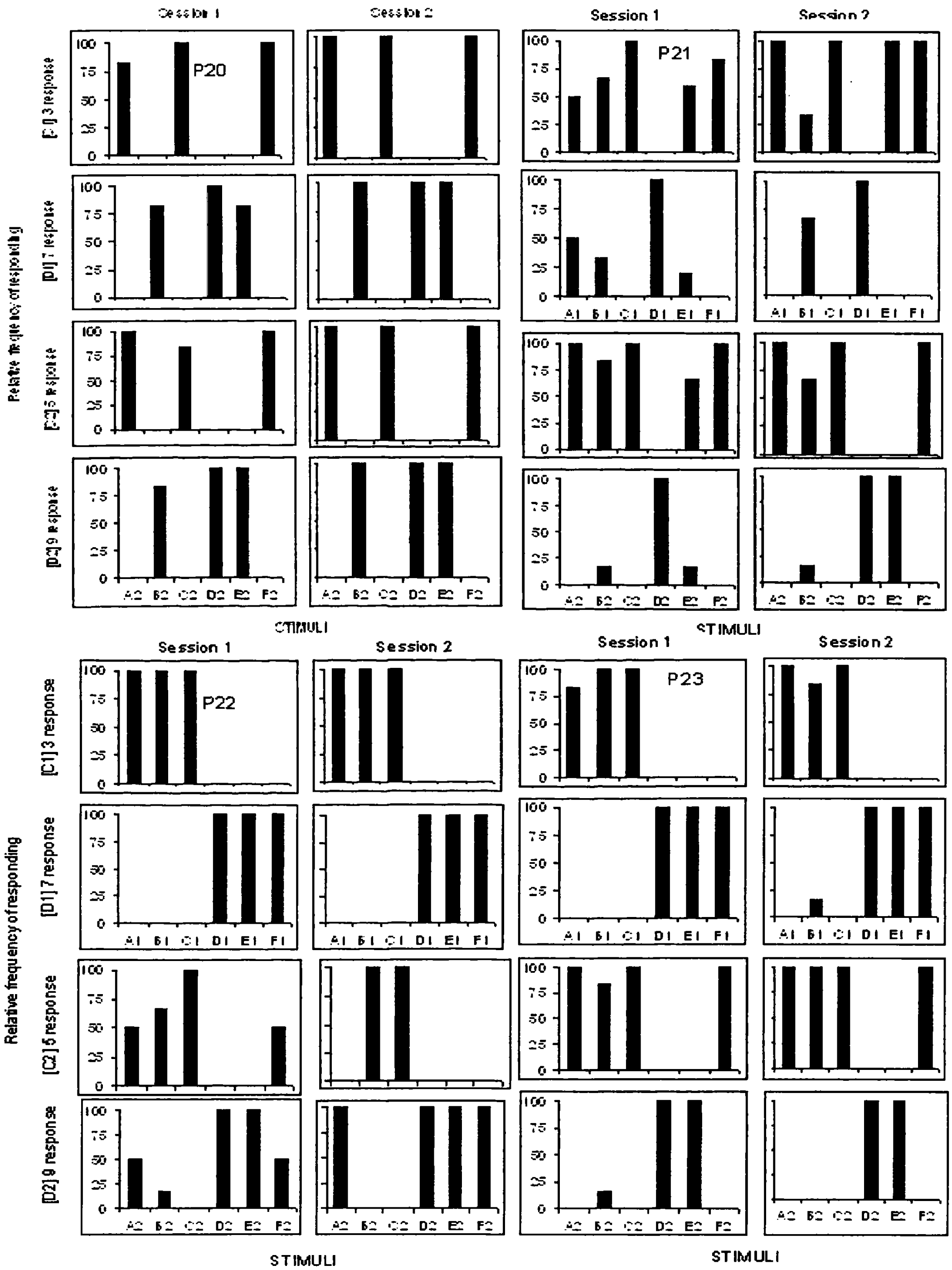


Figure 12: Relative frequency with which responses trained to C and D were evoked by stimuli in both classes across two sessions in the Equal Reinforcement group in Experiment 3

2.4.3 Discussion

The results from Experiment 3 suggested that level of reinforcement and training structures had no effect on the number of cycles participants needed to reach the mastery criterion on the simple discrimination training. The response patterns in the Unequal Reinforcement condition did not suggest that differential reinforcement and a serial training protocol is reflected in differential response transfer. Transfer from the D1 stimulus to the E1 and F1 stimuli was, as robust as the transfer from the C1 stimulus to the A1 and B1 stimuli. In the Equal Reinforcement condition, the untrained responses transferred from the C2 and D2 stimuli were under the control of nodal number, and that unequal reinforcement was not a prerequisite for nodal effects. In summary, delivering approximately an equal number of reinforcers for each trial type, and presenting all baseline trials in a simultaneous manner did not appear to have an effect on the probability of successful function transfer for any particular relation type. In addition, evidence of analysis of accuracy confirmed our findings in Experiment 1, that nodal number effect did not disappear under equal reinforcement, yet unlike Experiment 1, it became more generalized in Equal reinforcement condition, despite increased difficulty result from increased size of node number. In fact, all participants in Equal reinforcement group showed nodal effects in at least one block of their testing phase, and only one participant's responses in Unequal reinforcement group consistent with the prediction of nodal effect.

2.5 Experiment 4

The heterogeneity of response patterns in Experiment 3 suggested that there might be deficits in the training of the baseline relations. Experiment 4 was similar to Experiment 3 with the following exceptions. As the findings of Experiment 3 indicated that differential reinforcement did not reflect differential response transfer, an unequal reinforcement protocol was not employed in Experiment 4. Additionally, the criteria of 10 consecutively correct responses before proceeding to the test phase was expanded to 20 consecutively correct in order to stabilize performance.

2.5.1 Method

2.5.1.1 Participants

Fourteen participants began Experiment 4, (8 male; 6 female), ranging in age from 20 to 27 years (mean age = 23.14 years, standard deviation = 2.07 years). All participants were students (5 undergraduate, 9 postgraduate) at Swansea University and Swansea Institute. Participants were recruited through personal contacts by the first author (all Chinese nationals). All participants were naïve about the purpose of the experiment. The study was approved by the Department of Psychology, Swansea University, Ethics Committee.

2.5.1.2 Apparatus and Materials

The apparatus, setting, and stimuli were exactly the same as Experiment 3.

2.5.1.3 Procedure

The procedure was the same as Equal Reinforcement condition in Experiment 3, with the exception that 20 consecutively correct responses were required in the baseline training phase.

2.5.2 Results

Fourteen participants began Experiment 4, eight participants formed equivalence after repeated exposure to the training and testing phase.

The actual numbers of reinforcers delivered in Experiment 4 across all participants are presented in Table 16. These data indicate that the procedures employed were successful in manipulating the number of reinforcers across trial types that is, the number of delivered reinforcers was approximately equal across all trial types for all participants.

Table 16: Number of reinforcers delivered per training block per participant in Experiment 4.

Block	Trial Type																			
	AB	BC	CD	DE	EF	AB	BC	CD	DE	EF	AB	BC	CD	DE	EF	AB	BC	CD	DE	EF
	Participant 2					Participant 3					Participant 4					Participant 5				
1	20	19	19	18	19	45	45	45	44	45	86	86	85	84	84	24	22	22	23	22
2	8	8	8	8	9	15	16	16	16	16	11	11	11	11	10	38	39	39	39	39
3	6	7	6	6	6	10	10	10	11	10	8	8	8	8	8	4	4	4	4	4
4	10	10	11	10	10	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4					
6	7	7	7	7	7															
Total	55	55	55	53	55	78	79	79	79	79	113	113	112	111	110	70	69	69	70	69
	Participant 8					Participant 9					Participant 11					Participant 14				
1	26	25	24	25	25	62	62	62	63	62	12	12	12	12	10	59	59	59	60	59
2	8	7	7	7	7	22	21	21	21	21	12	12	11	12	12	5	5	5	6	4
3	15	14	14	15	15	4	4	4	4	4	5	5	5	6	6	4	4	4	4	4
4	7	6	6	6	7	6	7	6	6	6	4	4	4	4	4	5	5	5	6	5
5	5	4	3	4	4	4	4	4	4	4						8	8	8	6	8
6	5	4	4	4	4	4	4	4	4	4						4	4	4	4	4
7	4	4	4	4	4															
Total	70	64	62	65	66	102	102	101	102	101	33	33	32	34	32	85	85	85	86	84
	Participant 1					Participant 6					Participant 7					Participant 12				
1	22	21	20	21	20	56	57	57	57	57	36	36	38	37	38	78	78	79	78	79
2	4	2	2	4	2	3	4	4	4	4						6	6	7	7	6
3	6	8	7	7	8	14	14	12	13	13										
4	12	10	12	11	11	10	12	11	11	10										
Total	44	41	41	43	41	83	87	84	85	84	36	36	38	37	38	84	84	86	85	85
	Participant 13					Participant 15														
1	38	38	39	39	38	84	83	82	82	84										
2	4	4	4	4	4															
3	4	4	4	4	4															
Total	46	46	47	47	46	84	83	82	82	84										

2.5.2.1 Passed Participants

Table 17 shows the number of cycles needed for each participant to reach the 90% mastery criterion in the test phase. The variance reduced from previous experiment, again no one reached mastery criterion immediately after initial training.

Table 17: Number of cycles needed until participants reached the mastery criterion of 90% and above accuracy in class consistent responding in Experiment 4.

Participant No.	Cycles to Criterion
2	4
3	3
4	3*
5	2
8	5
9	4
11	2
14	3

(*) the first cycle of pt4 lack of test phase, due to the program failure

Table 18 depicts the percent correct on unreinforced baseline probes for each participant in test blocks. It indicated that performance on baseline conditional discriminations were improved by increasing the pass criterion from 10 to 20 consecutively correct trials.

Table 18: Percentage correct per test block for unreinforced baseline probes for passed participants.

Participant	Test Block Number						
	1	2	3	4	5	6	7
2	60	100	90	90	100	100	
3	90	90	100	100	100		
4	70	90	100	90			
5	90	80	100	100			
8	70	80	70	80	100	100	100
9	70	90	100	100	90	100	
11	100	90	100	100			
14	80	90	90	100	90	100	

2.5.2.1.1 Response Speed

Mean RSs were significantly different in terms of nodal numbers (L=216, P<.05). As shown in Figure 13, RS was an inverted function of nodal numbers, that is, as RS decreases, node number increases.

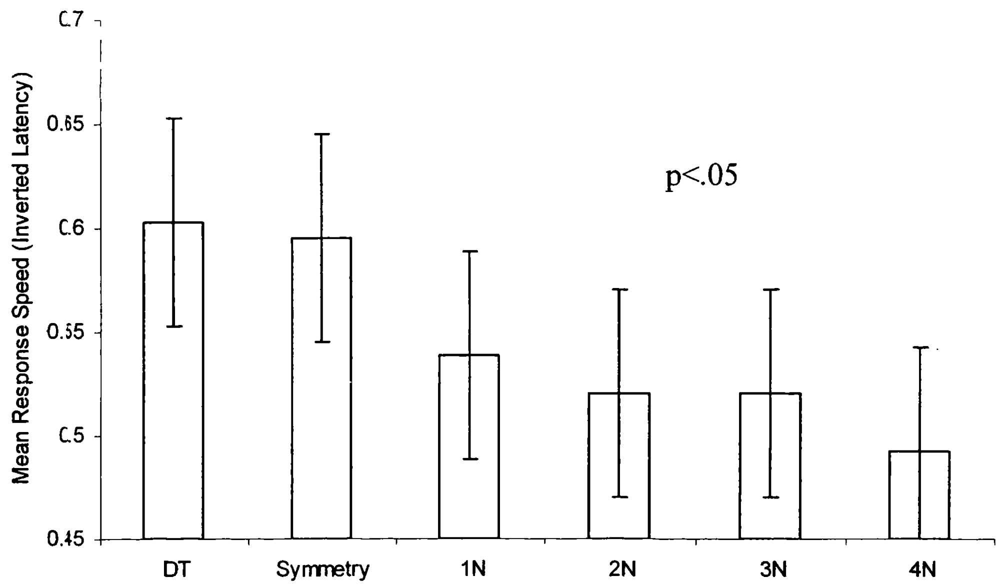


Figure 13: Mean response speed (Inverted latency) and 95% confidence intervals across all relation types for all participants in Experiment 4. DT: directly trained. 1N: 1 node. 2N: 2 nodes. 3N: 3 nodes. 4N: 4 nodes.

2.5.2.1.2 Response Accuracy

The top panel of Figure 14 shows the responses of Participant 2, 3, 4, and 5 in the test phase. After the initial testing block, a node effect was only recognized if 4-node trials were excluded for Participant 2 at Block 2 (1 node > 2 node > 3 node), a full node effect emerged at Block 3 and eventually reached ceiling at Block 4. Nodal effects were only recognized if 4-node trials were excluded for Participant 3 at Block 1 and 2 (1 node > 2 node > 3 node), responses reached ceiling at Block 3. The responses of Participant 4 showed a node effect at the first test block, and partially maintained at Block 2 (3node = 1 node > 2 node > 4 node). Responses reached ceiling at Block 3. Participant 5 quickly reached ceiling after the first block, and showed reversed order of node number effect when excluded 1-node trials (4 node > 3 node > 2 node).

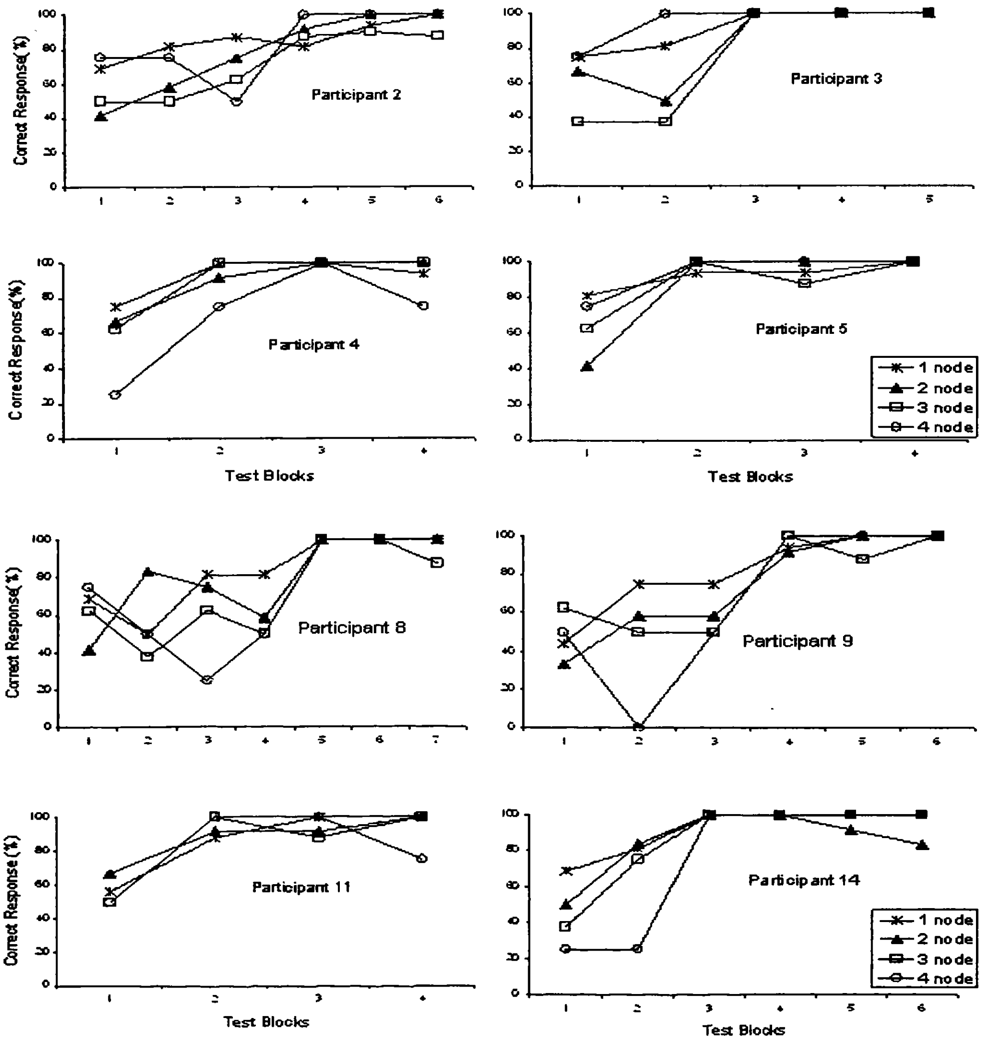


Figure 14: Percent correct for 1-, 2-, 3-, 4-nodes across blocks of testing in passed participants

The bottom panel of Figure 14 shows the responses of Participant 8, 9, 11 and 14 in the test phase. A reversed order of node effect was only recognized when 1-node trials were excluded from the first block for Participant 8 (4 node > 3 node > 2 node). It seems after self-correction, full node effect (1 node > 2 node > 3 node > 4 node) emerged at Block 3, partially maintained at Block 4 (1 node > 2 node > 3 node = 4 node), and reached ceiling at Block 5. After the initial testing block, Participant 9's responses showed full a node effect at Block 2, partially maintained at Block 3 (1 node >2 node > 3 node = 4 node), and reached ceiling at Block 4. Participant 14 showed node effect immediately after training, and this was partially maintained at the second block (1 node = 2 node > 3 node > 4 node), before performance reached ceiling at Block 3. Responses of Participant 11 reached ceiling shortly after training, therefore there was no evidence to base conclusions on for this participant.

2.5.2.1.3 Transfer Test

Response transfer was assessed by measuring the relative frequency with which the responses trained to the C and D stimuli were evoked by the other experimental stimuli. A series of exact tests were performed. The proportion of participants' responses that transferred to the A1 and B1 stimuli was significantly greater than those to transferred to the E1 and F1 stimuli, when trials were controlled by the C1 stimulus (p=.046, one-tailed); and was significantly greater to the E1 and F1 stimuli when trials were controlled by D1 stimulus (p=.006, one-tailed); when trials were controlled by the D2 stimulus, responses transferred to E2 and F2 stimuli approached significance, in contrast to those transferring to the A2 and B2 stimuli (p=.09, one-tailed). Relative frequency of responses for each participant after equivalence formation is shown in Figure 15.

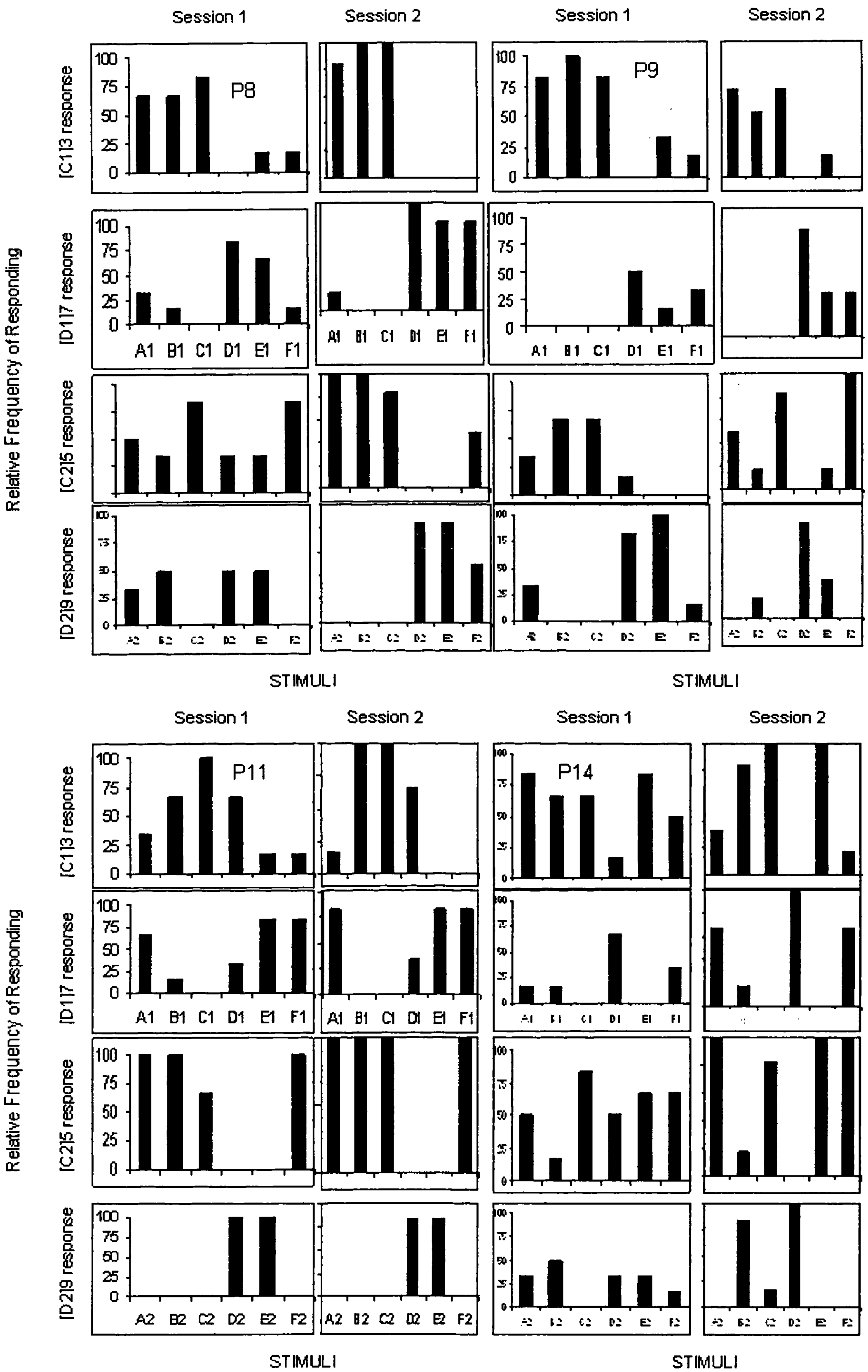


Figure 15: Relative frequency with which responses trained to C and D were evoked by stimuli in both classes across two sessions for passed participants

Responses of Participant 2 demonstrate a weak nodal effect in both Class 1 and 2 stimuli, with the A stimuli evoking the D, and C responses concurrently. The results of Participant 3 indicate some nodal effects across classes and sessions, with the F2 stimulus evoking more C2 responses than D2 responses, the A1 stimulus evoking the D1 response, and not the C1 response. Participant 4's responses do not appear to be under the control of nodal number, in spite of their C1 responses on Session 2. Participant 5's responses are similar, in spite of D1 and C2 responses on Session 1. The responses of Participant 8 appeared to be under the control of nodal number on Session 2, and not on Session 1. Participant 9's responses showed a nodal effect on D1 response across two sessions of response transfer tests. Participant 11's responses showed a weak nodal effect on Class 2 stimuli, with the F2 stimulus evoking the C2, and not the D2, response. The responses of Participant 14 do not appear to be under the control of nodal number, with responses distributed among members of the equivalence class.

2.5.2.2 Failed Participants

Four participants quitted Experiment 4 without reaching the mastery criterion of 90% correct in equivalence test, therefore no transfer tests followed. Performance on baseline probes in equivalence tests was shown in Table 19. Performance was less erratic in contrast to failed participants in Experiment 1 and 2.

Table 19: Percentage correct per test block for unreinforced baseline probes for failed participants

Participant	Test Block Number						
	1	2	3	4	5	6	7
2	60	100	90	90	100	100	
3	90	90	100	100	100		
4	70	90	100	90			
5	90	80	100	100			
8	70	80	70	80	100	100	100
9	70	90	100	100	90	100	
11	100	90	100	100			
14	80	90	90	100	90	100	

2.5.2.2.1 Response Accuracy

Percent correct responses for all nodal relations across all testing blocks were shown in Figure 16. Performance was generally erotic across participants, no consistent response patterns observed.

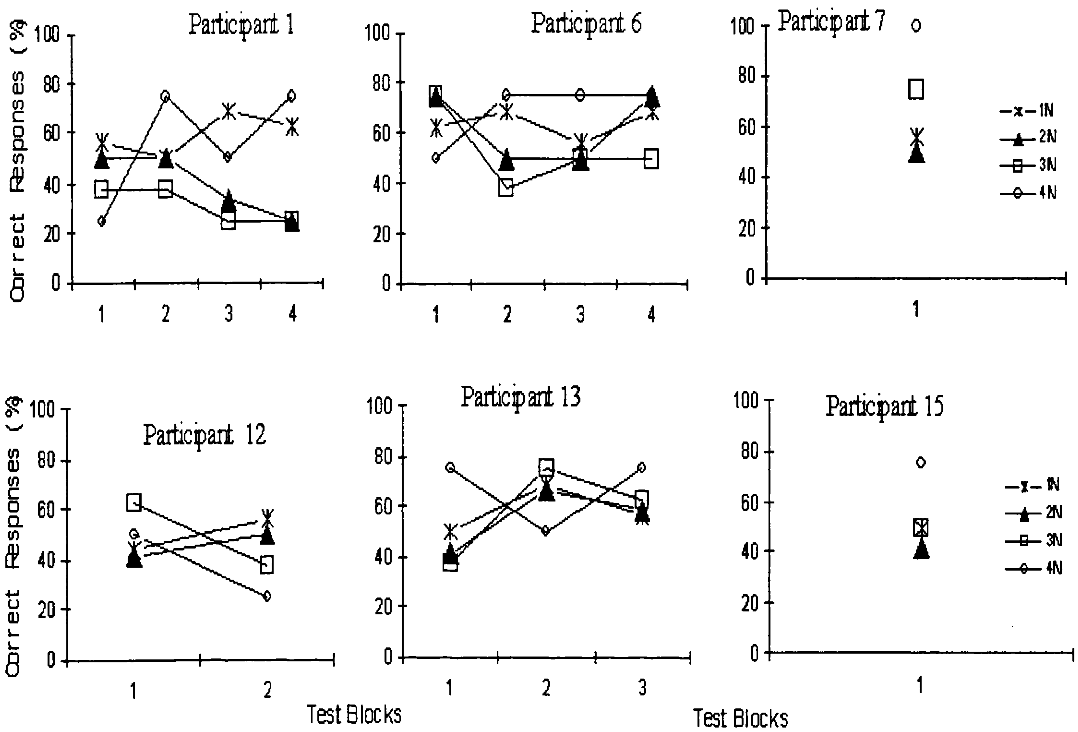


Figure 16: Percent correct for 1-, 2-, 3-, and 4-nodes across blocks of testing failed participants.

2.5.3 Discussion

Experiment 4 replicated the findings for the equal reinforcement group from Experiment 1 and 3, suggesting that, nodal effects remain intact under equal reinforcement and simultaneous training in the measure of RSs (Exp1) and transfer of function (Exp3). The strength of node effects obtained in Experiment 4 increased in both measures after stabilized baseline simple discrimination was demonstrated during testing. However, response accuracy seems disrupted, this might be due to extending the class members from 5 (Exp1) to 6 (Exp4). Despite this difficulty, 7/8 participants' correct responses still showed nodal effects at least once in their testing phase.

2.6 General Discussion

The results of Imam (2001, 2003, 2006) apparently demonstrated that nodal number effects were a function of a particular training structure, which resulted in unequal reinforcement across trial types during baseline conditional discriminations. We noted, however, that the interpretation of these studies is impossible, due to the incorrect use of a repeated-measures ANOVA, and because a comparison was made between conditions that were not equal in terms of power. Therefore, the present study sought to provide a systematic analysis of two types of training structure, which differed in terms of reinforcement delivered for particular trial types. The data suggested that nodal number was a predictor of response speed (Experiment 1) even when reinforcement for responding to baseline trial types during conditional discrimination training was equalized. An analysis of response accuracy proved that the time-accuracy trade off occurs in *limited hold* conditions only, as nodal effects disappear under both Equal and Unequal limited hold conditions. Nodal number effects did not disappear under simultaneous training protocol when measuring response accuracy; however it was less powerful when compared with the serial training group. In Experiment 2, participants were recruited only for the Equal no limited hold condition in order to examine the effect of sample size. Interestingly, nodal effects disappeared in both analysis of RS and RA. The expression or non-expression of nodal effects might be due to variables, such as, insufficient training, the format of MTS test. Therefore, in Experiment 3, a response-transfer test was employed as a more precise measure, and again a serial training structure that resulted in unequal reinforcement did not appear to influence nodal number. That is, less reinforcement did not result in poorer transfer. In addition, nodal effects became more generalized under equal reinforcement (Experiment 3 and 4). One factor that might have been affecting nodal performance is the number of baseline training trials. Sufficient and moderate training was necessary for the formation of equivalence classes without being masked by a ceiling effect (Experiment 4). When analyzing individual participant's data, half of the participants who passed the equivalence test demonstrated derived functional transfer within class members. In summary, the data of the present study suggest that a serial training protocol that results in unequal reinforcement is not a prerequisite for a nodal number effect.

In general, the proportion of participants who eventually pass the test for derived equivalence relations – the *yield* – is typically less than 100%. The yield varies according to factors such as training structure, class size and previous equivalence training (Fields et al., 1997). However, there is an exception in the current study, all participants eventually formed equivalence classes after repeated exposure to training and testing cycles in Experiment 3, which might be accounted for by chance. In Experiment 1, 2 and 4, There was a higher yield in the Unequal Reinforcement conditions – 4/5 participants in the no LH condition and 4/4 participants in the LH condition. In the Equal Reinforcement conditions the yield was 3/4 in the LH and 24/45 in the no LH condition. The higher yield in the Unequal Reinforcement condition, in which trial types were introduced in a serial manner, is consistent with previous literature (Adams, Fields, & Verhave, 1993). Interestingly, the yield was proportionally higher when the LH was in effect – 7/8, than when it was not – 14/25). The analysis of data from failed participants across all experiments did not show nodal consistent patterns, and contradict to studies suggesting nodal effects preserved even without equivalence formation (Fields, et al, 1995; Randall & Remington 1999).

There were some differences between the present study and the previous studies that varied the training structure and quantity of reinforcement (Imam, 2001, 2003, 2006). In the present study, the arrangement of training and testing trials was not manipulated. This was because the aim of the present study was solely to examine the effect of reinforcement and training protocol on nodal number, and not the influence of other factors such as training and testing protocol. A between-participant, rather than within-participant, design (Experiment 1 and 3) was employed and thus confounds due to effects of previous equivalence class formation were eliminated (see Fields et al. 1997, for a treatment of the facilitative effect of previous equivalence class formation). In common with many other equivalence studies (Fields et al., 1993; Barnes-Holmes et al., 2005), the present study employed two, rather than three, comparison stimuli. The use of only two comparisons stimuli on every trial has been criticized because the negative, rather than the positive, comparison may be the controlling stimulus (Carrigan & Sidman, 1992; Imam, 2001; Johnson & Sidman, 1993). However, Saunders, Chaney, and Marquis (2005) reported that control by the negative comparison does not seem to occur in practice. Moreover, Boelens (2002) has provided a theoretical treatment arguing in favor of two-option preparations.

The effects of nodal number have been shown to diminish with the addition of a third stimulus group (Kennedy, 1991 Exp. 2), and thus the inclusion of a third stimulus may have made interpretation difficult if nodal effects were not observed. Although studies have demonstrated Nodality effects with various types of stimuli (pictures, letters, symbols), the present study employed pronounceable letter strings, whereas Imam (2001, 2003, 2006) used graphical stimuli. Is it possible, indeed likely, that relations among stimuli were rapidly learned because pronounceable, rather than graphical, stimuli were employed (cf. Imam, 2001, 2003, 2006). However, the use of these stimuli did not seem to facilitate learning to such a degree that the differential reinforcement delivered between Unequal and Equal groups was attenuated, rather the training structure *per se* and numbers of baseline training counts. In Experiments 1 and 2 of the present study two 5-member equivalence classes were established, whereas in Experiment 3 and 4 two 6-member equivalence classes were trained and tested. In Experiment 1 nodal effects, were apparent in response speed during equal reinforced baseline training without a *limited hold*, which is consistent with recent finding (Fields, & Watanabe-Rose, 2008). In Experiment 2, nodal effects disappeared in both analysis of RT and RA. In Experiment 3 and 4, nodal effects became more generalized when reinforcement was held equal. In contrast, the use of a response-transfer test appeared to be as robust as accuracy score in detecting nodal number effects in both Experiments 3 and 4. An important consideration therefore, when designing studies to examine the effects of nodal number is that factors such as class size and task format interaction, that is, the number of baseline training trials may have an affect on the nodal number effects that are observed in some situations and not in others.

In general, the present experiments employed a simpler training and testing protocol than some previous equivalence studies. A pretraining procedure, such as matching uppercase and lowercase letters (Imam 2001) before the arbitrary conditional discrimination training was not employed, nor was a consequence-fading procedure (Fields et al., 1995). In the present study, the introduction of test trials was not signaled (c.f. Imam, 2001). This appears to have disrupted the performance in the test of response accuracy during the initial test blocks. Perhaps nodal effects would have been observed earlier in the test of response accuracy, and more consistently across test blocks if such a consequence-fading procedure were in place. A possible consequence of this minimalist approach was poor stimulus control by nodal number for some participants, but not others. However, the intention of the present study was to examine factors that influence delayed emergence and to examine if nodal effects would be *eliminated* under equal reinforcement and simultaneous training procedure. Therefore, it was desirable to avoid a ceiling effect (i.e., response accuracy equal to the mastery criterion immediately upon exposure to the test phase) as this would have made interpretation of nodal number effects difficult.

Imam (2001, 2003, 2006) and Spencer and Chase (1996), included only those trial types involving either the most trained conditional relations or the least trained conditional relations in their analyses. According to Spencer and Chase (1996), this exclusion procedure was designed partly to account for imbalances in "the order and differential amount of training on each baseline conditional discrimination" (p. 649). However, the aim of the present study was to measure the influence of these factors, and therefore it seemed counter-intuitive to exclude these trial types. Saunders and Green (1999, p. 132) have also noted that only including the most and least trained stimuli in the analyses introduces some additional problems in interpretation.

The present study investigated the roles of reinforcement and training protocol. Other possible confounds were not eliminated, however. For example, Saunders and Green (1999) suggested that previously untrained simple discriminations develop over the course of testing as a function of differential exposure to particular stimuli on unreinforced baseline trials during testing. These authors argued that the gradual acquisition of simple discriminations may result in response patterns that mimic nodal number effects.

In conclusion, the data from the present study indicate that nodal number effects are preserved, even when reinforcement is kept equal. These data therefore contradict those studies (Imam, 2001, 2003, 2006) that suggested nodal number effects were a product of unequal reinforcement.

Chapter 3: A Comparison of the Nodal and Discrimination Accounts of Equivalence Class Formation

3.1 Introduction

An equivalence class consists of a group of stimuli that have all become interrelated in spite of the fact that they do not necessarily share any physical properties in common with one another (Sidman & Tailby, 1982). In order to create equivalence classes subjects typically first learn a minimal number of relations between individual stimuli in a group. For example, if the group of stimuli consisted of the stimuli A, B, C and D, an equivalence class could be established by training two-term relations between AB, BC and CD stimuli using a conditional discrimination paradigm. Once a class has been established many new emergent relations form between the stimuli that were not previously taught.

Four types of emergent relations have been outlined by Sidman and Tailby (1982): (i) reflexive relations: where given the sample stimulus A subjects must be able to select A from an array of comparison stimuli, (ii) symmetrical relations: choosing A in the presence of B, B in the presence of C and C in the presence of D, (iii) transitive relations (A→C, B→D and A→D) (iv) equivalence relations (D→B, C→A and D→A)(Bush, Sidman & de Rose, 1989; Fields & Verhave, 1987). If all the emergent relations control responding, the group of stimuli functions as an equivalence class (Sidman, Kirk & Wilson-Morris, 1985) and the stimuli are substitutable for one another (Sidman, 1990, 1994).

The emergent relations formed after training have been described in terms of their nodal distance. A node is described as a stimulus linked by training to at least two other stimuli in a potential equivalence class (Fields, Verhave & Fath, 1984). For example, the 4- member class described above (A, B, C and D) contains two nodes (B and C). Of the transitive and equivalence relations formed, four (A→C, C→A, B→D and D→B) are comprised of stimuli separated by one node (B for the first two relations and C for the second). The remaining two relations (A→D and D→A) contain stimuli separated by two nodes (B and C).

In the equivalence literature, predicted outcomes often require repeated exposure to testing, this phenomenon is known as "delayed emergence". One possible explanation for the delayed emergence of relation types in equivalence formation has been addressed by the *nodal account* (Fields et al., 1984; 1987; 1993; 1995; Sidman, 1994). According to the nodal account, the gradual emergence of equivalence classes is a function of the nodal number that exerts differential control on class members, thus, response accuracy decreases and response time increases as a function of nodal number increases.

Saunders and Green (1999) argued that the frequency of baseline stimuli presented during testing results in the delayed emergence of relation types in equivalence. For example, if testing two five-member equivalence classes (AB, BC, CD, DE) with a simultaneous match-to-sample (MTS) training protocol, in the typical testing session baseline relations are re-entered along with novel/derived relations. Participants are instructed to choose one of two B stimuli comparisons (from different classes) in the presence of an A stimulus; to choose one of two C stimuli comparisons (from different classes) in the presence of a B stimulus; to choose one of two D stimuli comparisons (from different classes) in the presence of a C stimulus; to choose one of two E stimulus comparisons (from different classes) in the presence of a D stimulus. Therefore, the B, C, and D stimuli appear three times more often than the A stimulus, and approximately 50% more often than the E stimulus. Based on this observation, Saunders and Green (1999) predicted that due to the unbalanced frequency of stimulus presentation in baseline relations during the testing session, BD and DB relations will exert the most control in a relational discrimination, therefore, emerge first ("Stage1" Saunders and Green 1999, "SG 1" is used throughout the thesis), AE and EA relations will exert the least control, thus emerge last ("Stage3" Saunders and Green 1999, "SG 3" is used throughout the thesis), and the rest of relations (i.e., one-nodal relations for AC, CA, CE and EC and two-nodal relations for AD, DA, BE, and EB) will exert a secondary level of control in a relational discrimination, and, emerge second ("Stage 2" Saunders and Green 1999, "SG 2" is used throughout the thesis). This set of predictions is known as the discrimination account.

The nodal account of delayed emergence has received a great deal of attention from equivalence researchers (Fields & Verhave, 1987; Bentall, Jones, & Dickins, 1998; Fields, Landon-Jimenez, Buffington, & Adams, 1995; Imam, 2001; Kennedy, 1991; Kennedy, Itkonen, & Lindquist, 1994; Spencer & Chase, 1996). However, to our knowledge, no study to date has directly evaluated the discrimination account on equivalence performance (Saunders & Green 1999). The aim of Experiment 3.1 is to compare and contrast the predictions of both the discrimination account suggested by Saunders and Green (1999), and the nodal account, in order to determine which set of predictions can better account for delayed emergence of relation types in equivalence class formation. Other variables that may contribute to delayed emergence were kept constant.

3.2 Experiment 5

The current experiment trained and tested participants across two classes of five stimulus members on an MTS procedure with a simultaneous training protocol resulting in equal trial presentation on each trial type. A gradual feedback fading procedure was employed with feedback fading from 100%, followed by 50%, to 0% across training trials, as participants' performance can be impaired by the sudden elimination of feedback in the test phase. In order to stabilize performance, 16 consecutively correct responses across training trials with 0% feedback was required before proceeding to the equivalence test phase.

In a five member equivalence class, for example, according to the nodal account equivalence class formation emerges as a function of nodal number, that is, 1-node relations (AC, CA, BD, DB, CE, EC) emerge first, followed with 2-node relations (AD, DA, BE, EB), with 3-node relations (AE, and EA) emerging last. According to the discrimination account the formation of equivalence is a function of discriminations acquired during baseline training; thus, SG 1 relations (BD, and DB) should emerge first, followed by SG 2 relations (AC, CA, CE, EC, AD, DA, BE, and EB), while SG 3 relations (AE and EA) emerge last (see Table 19). As both accounts predict that AE and EA trials should emerge last, a comparison of these relations does not add to the understanding of the relative power of the nodal versus discrimination account, therefore it is not included in the current analysis. The following acronyms will be employed throughout the thesis when referring to different relations proposed in both discrimination and nodal accounts, respectively: SG1/1N (Stage 1 also 1-node relations) trial type, SG2/1N (Stage 2 also 1-node relations) trial type, SG2/2N (Stage 2 also 2-node relations) trial type.

If the nodal account is correct, responses to SG2/1N relation type should be faster and more accurate than responses to SG2/2N relation type. If the discrimination account is correct, responses to SG1/1N relation type should be faster and more accurate than responses to the SG2/1N relation type.

3.2.1 Method

3.2.1.1 Participants

Forty-seven healthy adults participated in the current experiment (18 male; 29 female), ranging in age from 18 to 34 years (m = 22.11 years, SD = 3.07 years). All participants were students (35 undergraduate, 12 postgraduate) from Swansea University and Swansea Institute; all participants were recruited through personal contacts by the first author (30 Chinese nationals, 9 British nationals, 4 Greek nationals, 3 Singapore nationals, and 1 Polish national), all were naïve about the purpose of the experiment, and were assured that they could withdraw without penalty at any time during the experiment. The study was approved by the Department of Psychology, Swansea University Ethics Committee.

3.2.1.2 Apparatus and Materials

The apparatus and materials were identical to Experiment 1- 4.

3.2.1.3 Procedure

A short questionnaire was attached to the consent form to record participants' age and gender. Each participant was also instructed to read and sign a consent form before starting the study. Participants were exposed to the experimental procedure individually; each cycle of the experimental procedure lasted between 20-35 minutes depending on participant's performance, and was scheduled within one day.

At the start of the experiment two equivalence classes were established by training AB, BC, CD and DE relations on a simultaneous basis, so that each trial type was presented the same number of times. The following instructions appeared on the computer monitor at the beginning of training:

> "In a moment three words will appear on the screen. Choose 1 of the 2 words in the lower corner of the screen by pressing the Z key for the left word and the M key for the right word. During some stages of the experiment, the computer will NOT tell you if your choices are correct or wrong. However, based on

what you have learned so far, you can get all of the tasks correct. Please do your very best to get everything right. Thank you and good luck."

Each trial started with the presentation of three stimuli, one at the top-centre (sample stimulus) and two at the bottom-corners of the screen (comparisons). Participants pressed the 'Z' or 'M' key on the computer keyboard to select the comparison on the left or right respectively.

A gradual feedback fading procedure was employed. Three phases were identified from the training session where feedback was presented after every trial in the first phase, hence 100% feedback until participants reached 16 consecutively correct responses; 50% feedback was then provided in the second phase until participants reached 16 consecutively correct responses. The training session was terminated when participants reached 16 consecutively correct responses without any feedback prompts, with the exception that, during the first 3 cycles for participant 4, and the first cycle for participant 5 only 8 consecutively correct responses were required in each feedback fading phase due to a programming error. Feedback when provided was in the form of the words "Correct" or "Wrong" appearing on the screen in red for 1-s, followed by 1.5-s inter-trial interval (ITI), during which the screen was blank.

Once the training session was completed, the test phase commenced without warning. All baseline conditional relations and tests for symmetry, transitivity and equivalence were presented in a single randomized block. Each type of trial was presented only once with 40 trials in total (see Table 20). Delayed emergence often requires repeated exposure to testing, therefore, the total number of each trial type participants were exposed to increase differentially depending on individual performance.

Table 20: Trial types per relation type.

Relation Type	Trial type			
Directly Trained				
	A1B1	B1C1	C1D1	D1E1
	A2B2	B2C2	C2D2	D2E2
Symmetry				
	B1A1	C1B1	D1C1	E1D1
	B2A2	C2B2	D2C2	E2D2
Stage 1/1 Node				
	B1D1	D1B1		
	B2D2	D2B2		
Stage 2/1 Node				
	A1C1	C1E1	C1A1	E1C1
	A2C2	C2E2	C2A2	E2C2
Stage 2/2 Node				
	A1D1	B1E1	D1A1	B1E1
	A2D2	B2E2	D2A2	B2E2
Stage 3/3 Node				
	A1E1	E1A1		
	A2E2	E2A2		

Each cycle consisted of training and testing. Participants were exposed to a minimum of 1 cycle until they reached the mastery criterion of 85% + accuracy on the test phase.

3.2.2 Results

The number of baseline relations that each participant was exposed to and their Standard Deviations (SD) are presented in Table 21. One participant quit the experiment without completing a single training session, therefore, their data is not included in the following analysis. The data that is summarised in Table 21 suggests that the control of equal delivery of baseline relation was successful (SD ranged from 0 to 1.5), and the actual number of trial presentations varied across participants.

Table 21: Number of baseline relations delivered in training and Standard Deviations (SD) across baseline relations for each participant.

	Baseline Relations															
	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE
	Participant 1				Participant 2				Participant 3				Participant 4			
Total	129	129	130	131	39	40	38	38	101	102	100	101	236	233	234	235
SD	0.96				0.96				0.82				1.29			
	Participant 5				Participant 6				Participant 7				Participant 8			
Total	111	110	113	112	48	48	46	48	163	163	164	162	86	86	87	87
	Participant 9				Participant 10				Participant 11				Participant 12			
Total	73	72	73	72	62	61	61	62	165	165	165	166	73	73	72	75
SD	0.58				0.58				0.5				1.26			
	Participant 13				Participant 14				Participant 15				Participant 16			
Total	62	61	62	61	178	179	179	178	152	150	149	151	175	176	173	174
SD	0.58				0.58				1.29				1.29			
	Participant 17				Participant 18				Participant 19				Participant 20			
Total	104	103	105	105	62	62	62	62	102	103	102	102	195	196	196	197
SD	0.96				0				0.5				0.82			
	Participant 21				Participant 22				Participant 23				Participant 24			
Total	31	32	30	30	104	104	104	103	53	52	52	52	74	73	74	73
SD	0.96				0.5				0.5				0.58			
	Participant 25				Participant 26				Participant 27				Participant 28			
Total	96	98	97	97	89	87	87	88	214	213	214	214	62	64	63	64
SD	0.82				0.96				0.5				0.96			

	Baseline Relations															
	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE	AB	BC	CD	DE
	Participant 29				Participant 30				Participant 31				Participant 32			
Total	119	118	121	121	100	100	99	99	130	130	130	131	61	62	62	64
SD	1.5				058				0.5				1.26			
	Participant 33				Participant 34				Participant 35				Participant 36			
Total	37	36	37	36	128	128	127	127	52	50	50	52	248	247	247	248
SD	0.58				0.58				1.15				0.58			
	Participant 37				Participant 38				Participant 39				Participant 40			
Total	271	272	269	272	58	57	58	58	105	105	106	108	133	133	134	134
SD	1.41				0.5				1.41				0.58			
	Participant 41				Participant 42				Participant 43				Participant 44			
Total	171	170	169	168	63	64	64	64	206	205	206	206	104	101	104	102
SD	1.29				0.5				0.5				1.5			
	Participant 45				Participant 46											
Total	72	70	71	73	102	102	102	103								
SD	1.29				0.5											

3.2.2.1 Passed Participants

Twenty-one participants reached the mastery criterion (7 male; 14 female), ranging in age from 18 to 34 years (mean age = 22.38 years, standard deviation = 3.20). Table 22 shows the percentage of correct responses per test cycle for the unreinforced baseline relations and the cycle number on which each participant reached the mastery criterion of 85% correct in the equivalence test. 15/21 participants reached the mastery criterion after their first exposure to the test session, which gave a much higher yield (70% passers after first exposure to the test session) than traditional simultaneous paradigms suggesting that the gradual feedback fading procedure stabilized performance after termination of the training session.

Table 22: Percentage correct per test cycle for unreinforced baseline probes for each participant until they reached mastery criterion of 85% and above accuracy in class consistent responding.

	Test Cycle Number							
Subjects	1	2	3	4	5	6	7	8
1	100	100	75	100	87.5	87.5		
2	100							
3	87.5	100	100					
4	87.5	87.5	100	87.5	100			
5	62.5	87.5	100	100	100	100		
6	100	100						
7	75	100						
8	62.5	100						
9	100	100						
10	100	100						
11	87.5	62.5	100	100				
12	75	87.5	62.5	87.5	87.5			
13	100	100	100					
14	75	100	75	62.5	87.5	100	100	
15	87.5	100	100	100				
16	100	100	100	100	100	100	100	100
17	100	87.5	100					
18	87.5	87.5	100					
19	62.5	87.5	100					
20	100	100						
21	100							

3.2.2.1.1 Response Speed

Response speed (RS; inverted latency) is calculated as this minimizes variance due to long latencies, which are more likely to be due to processes other than those of interest (e.g., inattention). Wilcoxon signed-rank tests were conducted in order to determine whether there was any trend of relations in terms of RSs, response speed was significantly faster on the SG2/1N trial type than to SG2/2N trial type ($p<.05$ two-tailed). There was no significant difference in response speed between the SG1/1N and SG2/1N relation types.

3.2.2.1.2 Response Accuracy

The percentage of response accuracy for each participant across all relation types (SG1/1N, SG2/1N, SG2/2N) to reach the mastery criterion is shown in Figure 17.

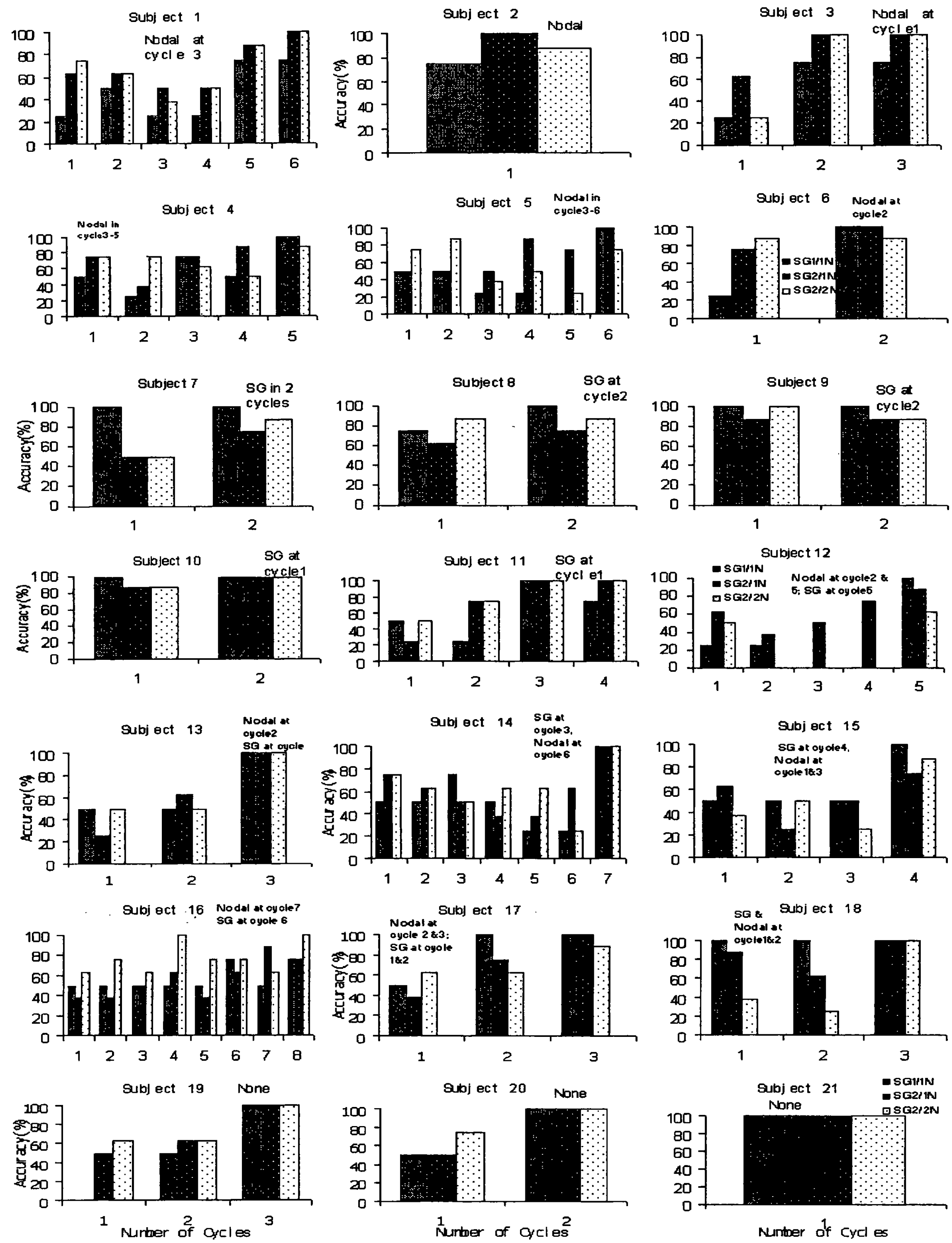


Figure 17: Percentage of response accuracy for each passed participant for the nodal account and discrimination account (SG) across test cycles.

By grouping condition consistent cycles (which is defined in the text in each bar chart) together across all testing cycles from all passed participants, the following summary was collated: Thirty per cent of cycles across all passed participants were consistent with the nodal account; 20 % of cycles across all passed participants were consistent with the discrimination account; 5% of cycles across all passed participants were consistent with both accounts. Participants reached ceiling performances on all three relation types in 11% of cycles. The remainder of cycles were inconsistent with either account. Moreover, 29% of passed participants showed sole consistency with the nodal account, 24% were solely consistent with the discrimination account, and 33% of passed participants demonstrated features of both accounts in different cycles, 14 % of passed participants showed features of neither account.

3.2.2.2 Failed Participants

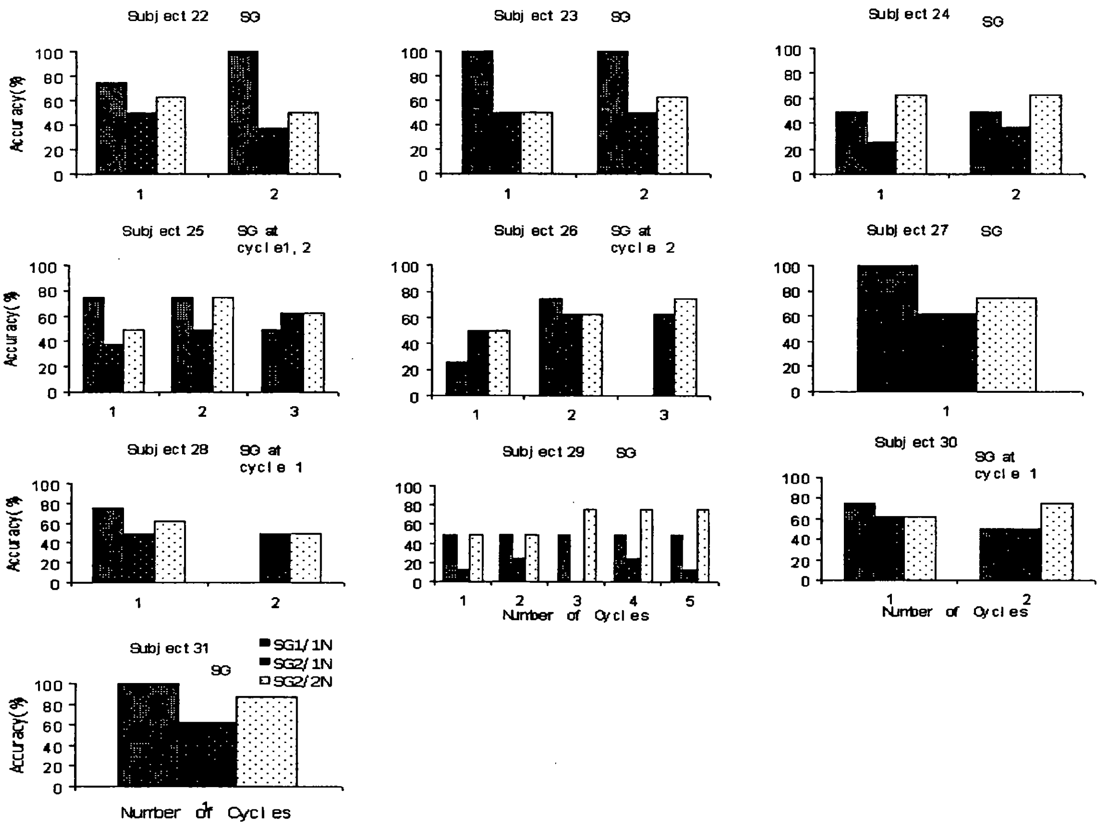
Eighteen participants quit the experiment before reaching to the mastery criterion, seven participants were mistakenly only tested to 70% accuracy, therefore all their data was analyzed in the failed participant group.

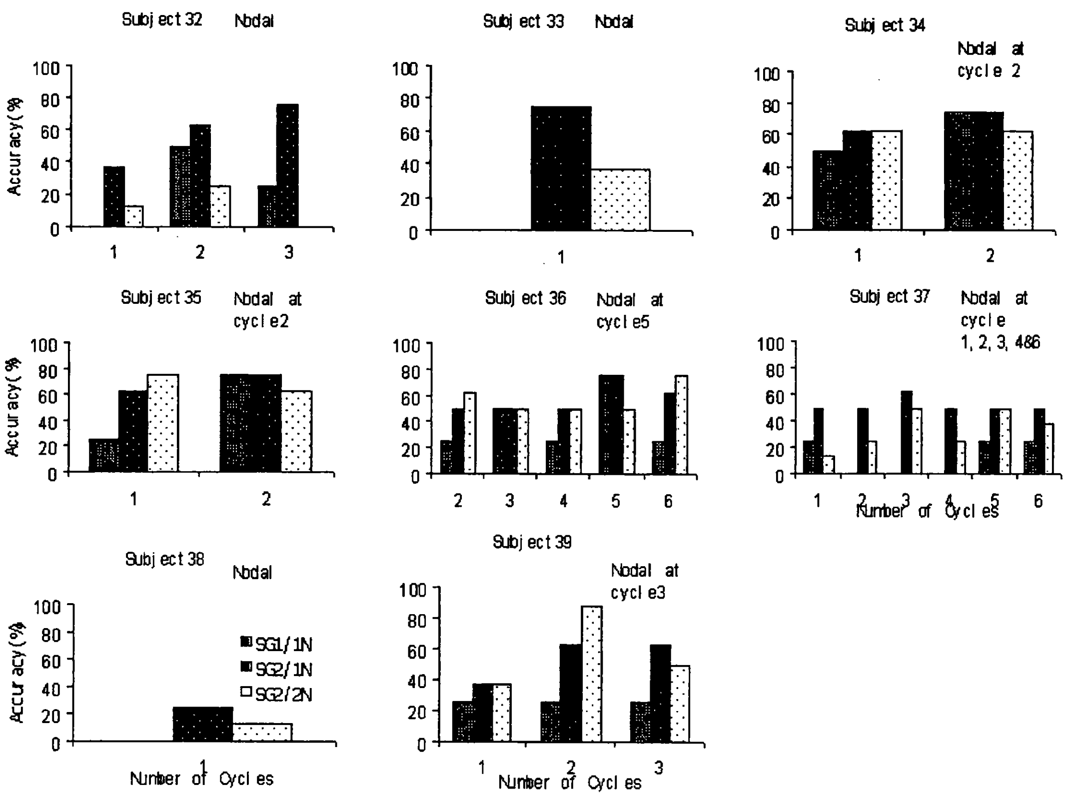
3.2.2.2.1 Response Speed

The same analysis was applied to failed participants as to those who passed. No significance was found across any of the comparisons.

3.2.2.2.2 Response Accuracy

Again, the same summary analysis was applied to failed participants as was applied to the participants who passed the mastery criterion in the test phase. The percentage of response accuracy for each failed participant across all relation types (SG1/1N, SG2/1N, SG2/2N) is shown in Figure 18.


Subject 22 SG
Subject 23 SG
Subject 24 SG
Subject 25 SG at cycle1, 2
Subject 26 SG at cycle 2
Subject 27 SG
Subject 28 SG at cycle 1
Subject 29 SG
Subject 30 SG at cycle 1
Subject 31 SG
SG1/1N
SG2/1N
SG2/2N
Accuracy(%)
Number of Cycles


Subject 32 Nodal
Subject 33 Nodal
Subject 34 Nodal at cycle 2
Subject 35 Nodal at cycle2
Subject 36 Nodal at cycle5
Subject 37 Nodal at cycle 1, 2, 3, 4&6
Subject 38 Nodal
Subject 39 Nodal at cycle3
SG1/1N
SG2/1N
SG2/2N
Accuracy(%)
Number of Cycles

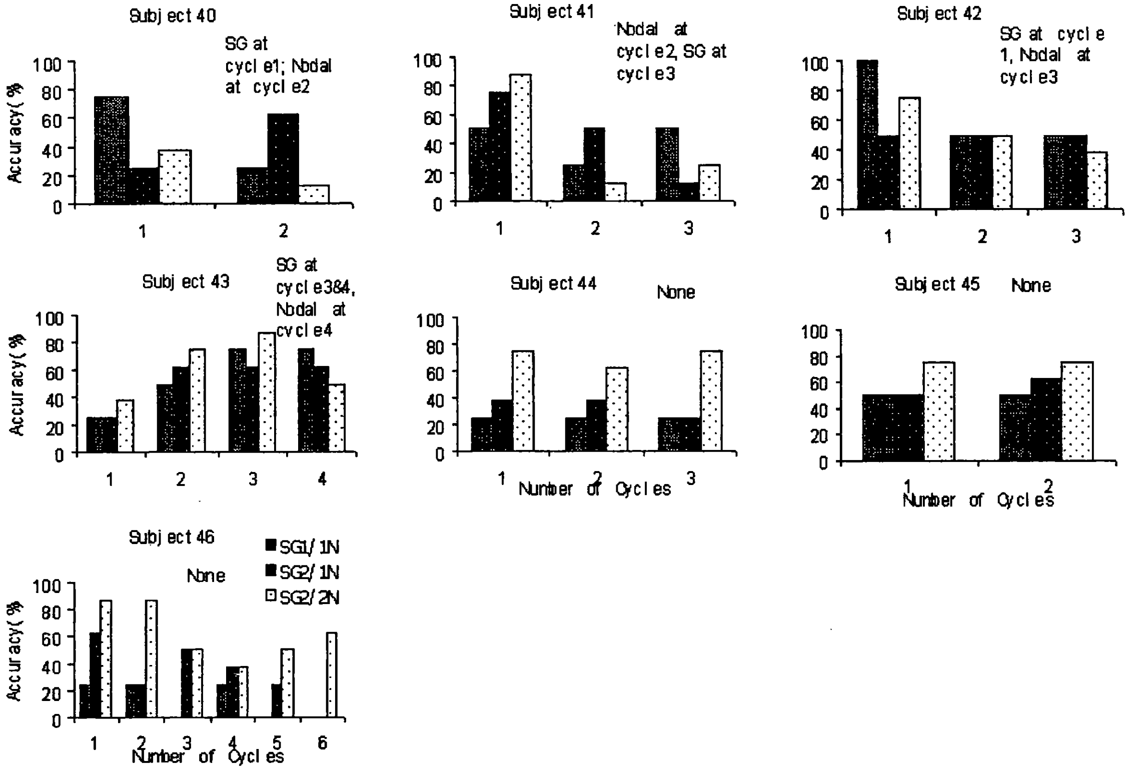


Figure 18: Percentage of response accuracy for each failed participant for the nodal account and discrimination account (SG) across test cycles.

To summarise Figure 18: Thirty-three per cent of cycles across all failed participants were consistent with the discrimination account; 26 % of cycles across all failed participants were consistent with the nodal account; only 1% of cycles across all failed participants were consistent with both accounts. The remainder of cycles were inconsistent with either account. Moreover, 40% of the failed participants showed sole consistency with the discrimination account, 32% were solely consistent with the nodal account, and 16% of participants demonstrated features of both accounts in different cycles, 12% of participants showed features of neither account.

3.2.3 Discussion

The main findings from Experiment 5 indicated that averaged RSs were significantly faster when responding to the SG2/1N trial type than to the SG2/2N trial type ($p<.05$). However, no difference was found between discrimination account relations (i.e., SG1/1N vs. SG2/1N). Correct responses across participants who formed equivalence after repeated exposure to training and testing cycles again favoured the nodal account, as there were more nodal consistent cycles than discrimination consistent cycles (22 vs. 15). There is little to suggest that the discrimination account is an accurate predictor of delayed emergence, even though there is a greater consistency to the discrimination account in correct responses from people who did not formed equivalence classes.

Overall the results from the current experiment lend support to the prediction that delayed emergence is not the result of discriminations formed during baseline training, as suggested by Saunders and Green (1999). Rather, an engram (Semon, 1921) plays a certain role in the formation process, particularly in the case of response speed. This assumption is in line with research conducted by Fields and his colleagues' over the last two decades, who have tried to tease out nodal effects in a variety of equivalence paradigms, across a wide range of behavioural and neurological data, even in the absence of equivalence formation (Fields, et al, 1995; Randall & Remington, 1999). These researchers have argued that test performance is contingent upon reinforcement history, and that the test format emphasis on either between-class discrimination in different class sets or class-based relations (e.g. nodality) in the same class was responsible for the expression of nodal effects in some studies, but not others (see Fields and Moss 2007 for a systematic review).

According to Fields and Moss's (2007), the current test format emphasised the between-class discrimination in different class sets, which theoretically maximized the between-class discrimination control, this might account for the absence of nodality in some test cycles for some participants. However, even in light of this difficulty, more than half (13/21) of the participants who were successfully trained in equivalence (after repeated exposure to the training and testing cycles) demonstrated nodal patterns in their correct responses on certain cycles. This again, supported the reinforcement contingency theory of equivalence formation (Sidman, 1994) expanded by Fields and Moss 2007, that is, the conditional discrimination acquired during training is composed of both class-consistent function between class sets and a nodal structure function within a class.

Aforementioned, the current experiment was the first to attempt to compare the nodal account with the discrimination account in explaining the delayed emergence of relation types in equivalence class formation. The analysis of correct responses outlined in the methods section provided a good example of how to approach the two theories (nodal and discrimination accounts) with the same data set. One might argue that the nature of comparisons is not mutually exclusive, responses likely fall into a linear pattern in which participants demonstrate better responding on SG1/1N compared to SG2/1N, followed by SG2/2N relations. However, this was not the case in the current study as participants demonstrated such a linear pattern on only 4/74 of the cycles in passed participants, and 1/69 of the cycles in failed participants, which is way beyond chance level.

Two equivalence classes were trained and tested in the current study, to provide a sample analysis of delayed emergence of relation types in equivalence. However, increasing the number of equivalence classes which result in increased magnitude of power might provide more interesting results (e.g., three five-member equivalence classes result in *four* times more presentation of stimuli B, C, and D, than A stimulus; and 75% more than E stimulus). In addition, a simple MTS training and testing paradigm was employed in this study, and only response time and accuracy were recorded. Future research might attempt to include a function-transfer paradigm as a more sophisticated measure (as in Fields et al., 1996).

To summarise, the current findings provided evidence for a nodal structure acquired contingent upon reinforcement history which is in line with Fields and Moss's (2007) expanded analysis of contingency theory, that is, the conditional discrimination acquired during training is composed of not only a class-consistent function between class sets but also a nodal structure function within a class.

Chapter 4: The Role of Reinforcement, Number of Stimulus Presentations, and Time of Acquisition on Equivalence Class Formation/Semantic Network Growth

4.1 General Introduction

Semantic network theory attempts to describe how language is organized in the human brain (Steyver & Tenenbaum, 2005). A semantic network is composed of words or knowledge (i.e. nodes) and associations or relations (i.e. connections). A semantic network represents the interaction of a number of words through direct and indirect relations in a person's language repertoire and it provides a structured representation of language acquisition and inference (Sidman, 1994). Recently there has been increasing evidence demonstrating that traditional network theories of verbal or semantic meaning (Anderson, 1976, 1983; Collins & Loftus, 1975; McClelland & Rumelhart, 1988) share similarities with derived stimulus relations (Hayes & Bisset, 1998; Barnes & Hampson, 1993; Cullinan, Barnes, Hampson, & Lyddy, 1994; Fields & Verhave, 1987; Barnes-Holmes, et al, 2005; Whelan, et al, 2005). The literature on derived stimulus relations accounts for language acquisition in terms of control by discriminative stimuli during explicit training, which ultimately results in between-class discriminations and class-related discriminations (i.e., node effects) amongst stimuli. The simplest form of derived stimulus relation, the equivalence relation, can be viewed as the simplest composite unit of a semantic network (Sidman, 1994, Hayes, et al., 2001).

The most widely employed measure of relatedness among nodes in a semantic network is the semantic priming task (White, 1986). Priming tasks involve comparing participants response times to semantically related words (e.g. dog – cat) to unrelated words (e.g. dog – chair). Faster responding to words that are semantically related compared to unrelated is known as priming or the "priming effect". The most commonly employed experimental preparations in priming studies have been the lexical decision (White, 1986) and pronunciation tasks (Brysbaert, 1996). In a typical lexical decision task, word pairs (i.e., prime and target) are presented consecutively, for a short period of time. Participants are typically instructed to respond as to whether the words presented are English (e.g. wind) or not (e.g. inwd). While in the pronunciation task, participants are instructed to say each of the word presentations aloud. There are many variations of experimental parameters in these tasks, for example, they may involve single-word priming (i.e., responding only to the target) or prime-target onset asynchronies (SOAs) (i.e. measuring the latency between prime and target). Based on different types of prime-target relations, priming effects observed are classified as semantic (i.e.. semantically related), associative (i.e., associatively related), mediated (i.e.. indirect relation between the prime (e.g. lion) and target (e.g. stripe) via an associated other word (e.g. tiger), and episodic priming (i.e. the relation between the prime and target was learned earlier in the experimental session; see Neely, 1991 for a systematic review). Mediated priming most resembles derived stimulus relations as this type of priming is generated from word pairs that have no direct associations between them (e.g. lion, stripe) Therefore, behavioural studies of this kind typically focus on establishing mediated priming via equivalence class formation.

The first behavioural study that has sought to test the priming effect in a lexical decision task was conducted by Hayes and Bisset (1998). In their study, three 3-member equivalence classes comprising of nonsense words were trained and tested using a match-to-sample procedure. Subsequently, the participants were exposed to a two-word lexical decision task, in which they were instructed to press a "YES" key if both words were from the previous training and testing phase, and to press a "NO" key if one or both words were not seen before. Mean RTs were significantly faster on trials involving equivalently related word pairs than non equivalently related pairs, reflecting the mediated and episodic priming typically found in the semantic network literature.

However, there were few limitations in the Hayes and Bisset, (1998) study. First, the use of a two-word lexical decision task rather than the more commonly used one-word priming paradigm limits the generality of their findings. Second, feedback was provided during the lexical decision task, it is possible that priming may disappear without feedback, therefore the feedback may have confounded their findings. Third, and more importantly, the test for equivalence in their study was presented to participants immediately following training. Therefore, it was unclear whether the emergent lexical decision performances would have occurred in the absence of an equivalence test. A more subtle example of derived semantic priming (i.e., mediated priming) would be demonstrated in the absence of the pre lexical equivalence test. It is also possible that mediated priming would not happen without an equivalence test. Barnes and his colleagues (1995) found that participants produced a correct derived performance, following repeated failures, when they were exposed to an equivalence testing procedure. The authors suggested that this provided evidence that testing procedures may sometimes facilitate derived behaviour (see Barnes & Keenan, 1993, p.78). Specifically, equivalence testing procedures expose participants to trials that pair the derived C-A stimuli together. Thus, "priming...observed in Hayes and Bisset's study may have simply reflected direct rather than mediated priming" (Barnes-Holmes et al 2005, p 423).

In order to address the limitations of Hayes and Bisset (1998) Barnes-Holmes et al., (2005) conducted a series of three experiments. They extended on the work of Hayes and Bisset (1998) by integrating Event Related Potential (ERP) as an additional dependent measure of emergent priming (Exp3), this aspect of the work will be discussed in Chapters 5 (see Section 5.1 for more detail). Experiment 1 in the Barnes-Holmes et al., (2005) study was designed to address the first two limitations from Hayes and Bisset's (1998) study. To that end, two 4-member equivalence classes were trained and tested and this was followed by a single-word lexical decision task without any feedback. The Mean RTs replicated the findings from Hayes and Bisset's original study, that is, the mean RTs were significantly faster when both the prime and the target words were from the same equivalence class than when they were from different classes. Experiment 2 was designed to overcome the third limitation, specifically, the lexical decision task was presented prior to the equivalence test phase. Findings from their Experiment 2 revealed mediated priming in equivalently related words only in those participants who had subsequently passed the equivalence test. Thus, suggesting that the equivalence test may have a facilitative role in semantic learning. Again, Barnes-Holmes et al., (2005) supported the general findings from Hayes and Bisset (1998), which closely resemble those reported in the semantic network literature. Research in the area of derived stimulus relations in semantic network growth has provided a contingency based explanation consistent with the behavioural tradition (Sidman, 1971). Specifically, according to this account, the strength of connections between nodes is a function of the number of reinforcers with which they are associated.

Alternatively, some semantic network models have predicted a correlation between the time at which a word enters a network and the strength of the relations to that word (Turner, Valentine, & Ellis, 1998; Morrison & Ellis, 1995, 2000; Brysbaert, Lange, & Van Wijnendaele, 2000; Barry, Morrison & Ellis, 1995; Ellis & Morrison, 1998; Gerhand & Barry, 1999; Moore & Valentine, 1998; Lewis, 1999). This is known as "Age of Acquisition" (AoA), that is, participants respond faster to words that are learnt earlier. For example, Izura and Ellis (2002) found the AoA in lexical decision and pronunciation tasks in the acquisition of first and second languages in Spanish-English bilinguals. A third possible contributing factor in semantic network growth is known as familiarity or word frequency, that is, the more often a word is encountered, the more efficient the synaptic connections representing this word in the network become (Brysbaert, 1996; Gerhand & Barry, 1998; Morrison & Ellis, 2000). For example, Brysbaert and his colleagues (1996, 2000) demonstrated frequency has a clear effect on a lexical decision task in Dutch language participants. Although all of the above factors have been tested separately indicating that each process contributes independently to semantic learning, the interaction among the processes involved in forming and maintaining links among nodes has not been directly tested. For example, there has been debate on whether early AoA increases connection strength independently of word frequency, as high-frequency words are likely to be acquired earlier than low-frequency words (see Ghyselinck, et al., 2004, for a review). Similarly, there has been disagreement over whether frequency of encountering words or reinforcement for using those words results in stronger connections (Iman, 2001). The current study systematically manipulates these factors in order to examine the interaction among the three processes, which might shed some light on our understanding of semantic learning.

The current study replicated the behavioural paradigm employed by Barnes-Holmes et al (2005) in their Experiment 2, in which, two 4-member equivalence classes were trained using a MTS protocol, followed by a single-word lexical decision task with the test for equivalence after the lexical decision task. The uniqueness of the current study, however, is that is aims to systematically manipulate the level of reinforcement, frequency/number of stimulus presentations and AoA, to test their relative contribution to the process of semantic learning. Therefore, if semantic learning is indeed the result of these three key processes, it is predicted that participants exposed to 100% reinforcement, unequal number of presentations, and phased acquisition will demonstrate the strongest connections (i.e. shortest RTs among the 100% reinforcement, unequal no. of presentations, and phased acquisition condition), while 50% reinforcement, equal number of presentations and simultaneous acquisition will demonstrate the weakest connections (i.e. longest RTs for the 50% reinforcement, equal no. of presentations and simultaneous acquisition condition) Also, reinforcement will be the most influential main factor when compared to AoA and no. of stimulus presentations.

4.2 Experiment 6

4.2.1 Method

4.2.1.1. Participants

249 healthy adults participated in the current experiment (84 male; 165 female), ranging in age from 18 to 60 years (m = 22.52 years, SD = 5.75 years). The participants were either students from Swansea University, Swansea Institute, or personal contacts of the experimenter. All of the participants were naïve about the purpose of the experiment, and were assured that they could withdraw without penalty at any time during the experiment. The study was approved by the Department of Psychology, Swansea University Ethics Committee.

4.2.1.2. Design

The current experiment employed a 2x2x2 betweens subject design, with reinforcement (100% vs. 50%), no. of stimulus presentations (Unequal vs. equal), and time of acquisition (phased vs. simultaneous) across conditions as the between subject factors (See Table 23 for details of each condition)

4.2.1.3 Apparatus and Materials

The apparatus and materials were identical to Experiment 1- 4, with the exception that two groups of semantically related English words (i.e. summer, winter, spring, autumn; nose, eye, finger, toe) were employed for practice purposes before baseline stimuli and novel stimuli (see Table 2)were introduced into the lexical decision task.

4.2.1.4 Procedure

A short questionnaire was attached to the consent form in order to record participants' age and gender. Each participant was also instructed to read and sign a consent form before commencing the study. Participants were exposed to the experimental procedure individually which lasted approximately 45-60 minutes which varied depending on individual performance.

Each participant completed three phases of computer generated tasks: 1). Match-to-Sample equivalence training, 2). A single-word Lexical Decision Task (LDT), and 3). Equivalence testing.

Phase1: Match-to-Sample Training. Six baseline relations were trained using a delayed match to sample procedure across 180 trials (see Table 23 for each trial type).

Table 23: The trial types presented during training

Directly Trained			
	A1B1	B1C1	C1D1
	A2B2	B2C2	C2D2

Each participant was presented with one sample stimulus (i.e., A1, B1, C1), and two comparison stimuli (i.e., B1B2, C1C2, D1D2), and then trained to select the class consistent stimulus from the two comparison stimuli. The on-screen instructions are shown below:

> "During this phase of the experiment, you will be trained to match FOREIGN WORDS to other FOREIGN WORDS. All words in this phase will be TRUE FOREIGN WORDS. Your task is to look at the foreign word at the top of the screen, and then look at the two foreign words at the bottom of the screen on the left and right when they appear. Your task is to match one of the two foreign words at the bottom of the screen to the foreign word that appeared at the top of the careen. You should choose one of these by pressing the Z or M key on the keyboard in front of you.
>
> Look at the word in the middle of the screen, then look at the two words on the left and right. Your should choose one of these two words by pressing the Z and M key on the keyboard in front of you.
>
> THE RELATION BETWEEN THE FOREIGN WORDS IS NOT ALREADY KNOWN TO YOU. YOU WILL HAVE TO LEARN BY

TRIAL AND ERROR. REMEMBER, YOUR TASK IS TO PICK THE FOREIGN WORD ON THE BOTTOM THAT GOES WITH THE ONE AT THE TOP."

Participants were randomly assigned to eight experimental conditions (see Table 24). All eight conditions started by establishing two four-member equivalence classes AB, BC, and CD relations with the systematic manipulation of reinforcement probability, time of acquisition and no. of stimulus presentations dependent on their assigned condition.

Table 24: The eight conditions and their abbreviations in term of reinforcement, frequency and time of acquisition.

Condition 1	100% Reinforcement Unequal no. of Presentations Phased Acquisition
Condition 2	100% Reinforcement Unequal no. of Presentations Simultaneous Acquisition
Condition 3	100% Reinforcement Equal no. of Presentations Phased Acquisition
Condition 4	100% Reinforcement Equal no. of Presentations Simultaneous Acquisition
Condition 5	50% Reinforcement Unequal no. of Presentations Phased Acquisition
Condition 6	50% Reinforcement Unequal no. of Presentations Simultaneous Acquisition
Condition 7	50% Reinforcement Equal no. of Presentations Phased Acquisition
Condition 8	50% Reinforcement Equal no. of Presentations Simultaneous Acquisition

In the equal no. of presentations conditions, trials from both equivalence classes (i.e., Class1 and Class2) appeared an equal number of times (i.e., 10 each), whereas in the unequal no. of presentations conditions trials from Class 1 appeared 7 times, trials from Class 2 appeared 3 times, thus examining if greater no. of stimulus presentations led to stronger connectedness among stimuli. For the 50% reinforcement conditions, all correct or incorrect responses were reinforced or punished ("Correct" or "Wrong" feedback) on 50% of trials, whereas in the 100% reinforcement conditions all responses were reinforced or punished across 100% of trials, thus examining if more reinforcement led to stronger connectedness. For the simultaneous acquisition conditions, all trials appeared simultaneously from the start, whereas for the phased acquisition condition all Class 1 trials were presented before Class 2 trials, thus examining if phased acquisition led to stronger connectedness. In the equal no. of presentation phased acquisition conditions, the first 60 trials were from Class 1 and subsequent trials were presented according to a 30: 90 ratio of Class 1: Class 2 to preclude recency effects. In the unequal no. of presentation phased acquisition conditions, the first 84 trials were from Class 1, and subsequent trials were presented according to a 42: 58 ratio of Class 1 : Class 2.

Phase 2: Lexical Decision Task. A single-word lexical decision task modelled on Barnes-Holmes, e al., (2005) was used to examine mediated priming effects. In this phase, baseline relations trained in phase 1 were presented in addition to novel pseudo-word control pairs. Participants were informed that the baseline stimuli were true foreign words in order to help them distinguish these stimuli from the novel pseudo-word pairs. Participants were asked to respond if they recognised the target (as presented before) or not. The on-screen instruction was shown below:

> "Now that you have had some practice, let's begin using FOREIGN and NONSENSE WORDS. During this phase of the experiment, you will be asked to respond to some words on the computer screen. SOME of these words will be FOREIGN words you have just learned. BUT some of the words will be NONSENSE words. Two words will appear on the screen, one after the other. You MUST observe the first word that appears and pronounce it mentally to yourself. When the second word appears your task will be to press the YES KEY if that SECOND word is a foreign word (that you were exposed to earlier) or the NO KEY if the SECOND word is not a foreign word. Remember you should respond YES or NO only to the SECOND word. You should observe and mentally read the first word, but your response should be to the SECOND word alone
> It is also VERY IMPORTANT that you respond as quickly and accurately as possible on every trial. The computer will be recording your response time and accuracy on every trial. "

Each trial began with a warning stimulus, the presentation of a red "X", in the middle of the screen. This X remained on the screen for 500 ms, and was then replaced by the prime (e.g. A1), which remained on screen for 200 ms. When the prime was removed from the screen there was a 400 ms blank, and then the target stimulus (e.g. B1) was presented (a stimulus onset asynchrony [SOA] of 600 ms). After 1,500 ms the target was removed and a green "X" appeared in the middle of the screen. Finally, after 1,250 ms the green X was replaced by the red X and the next trial began.

Two groups of semantically associated common English words and two groups of novel nonsense words were used as a brief practice phase before introducing to the block of randomly mixed equivalence stimulus pairs and non-equivalence pairs. During the lexical decision task, each participant was presented with 24 pairs of stimuli that were from the same equivalence relations, and 32 pairs that were from different equivalence relations and 102 trials that contained one or two previously unseen nonsense stimuli (e.g. A3, B3, C3, D3, A4, B4, C4, and D4). All nonsense words, including baseline and novel words, were presented in Table 2. Table 25 presents all 158 trial types during the lexical decision task.

Table 25: 158 trial types presented during the lexical decision procedure. Pm = Prime, Tg = Target, Rp = Correct Response, A3, B3, C3, D3, A4, B4, C4, and D4 = Nonsense Word.

Within Class			Cross Class			Class-nonsense			Nonsense-class			Nonsense-nonsense		
Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp
Directly Trained			A1	A2	Yes	A1	A4	No	A3	A1	Yes	A3	A4	No
A1	B1	Yes	A1	B2	Yes	A1	B4	No	A3	A2	Yes	A3	B4	No
B1	C1	Yes	A1	C2	Yes	A1	C4	No	A3	D1	Yes	A3	D3	No
C1	D1	Yes	A1	D2	Yes	A2	A3	No	A3	D2	Yes	A3	D3	No
A2	B2	Yes	B1	A2	Yes	A2	B3	No	A3	C1	Yes	A3	C4	No
B2	C2	Yes	B1	B2	Yes	A2	A4	No	A3	C2	Yes	A4	A3	No
C2	D2	Yes	B1	C2	Yes	A2	B4	No	B3	B1	Yes	A4	B3	No
Symmetry			B1	D2	Yes	A2	C4	No	B3	B2	Yes	A4	C4	No
B1	A1	Yes	C1	A2	Yes	B1	B3	No	B3	C1	Yes	A4	D4	No
C1	B1	Yes	C1	B2	Yes	B1	C3	No	B3	C2	Yes	A4	B4	No
D1	C1	Yes	C1	C2	Yes	B1	B4	No	B3	B1	Yes	B3	B4	No
B2	A2	Yes	C1	D2	Yes	B1	C4	No	B3	B2	Yes	B3	C4	No
C2	B2	Yes	D1	A2	Yes	B1	D4	No	C3	C1	Yes	B3	A3	No
D2	C2	Yes	D1	B2	Yes	B2	B3	No	C3	C2	Yes	B3	A3	No
Equivalence			D1	C2	Yes	B2	C3	No	C3	B1	Yes	B3	D4	No
A1	C1	Yes	D1	D2	Yes	B2	B4	No	C3	B2	Yes	B4	B3	No
B1	D1	Yes	A2	A1	Yes	B2	C4	No	C3	A1	Yes	B4	C3	No
A2	C2	Yes	A2	B1	Yes	B2	D4	No	C3	A2	Yes	B4	D4	No
B2	D2	Yes	A2	C1	Yes	C1	C3	No	D3	D1	Yes	B4	A4	No
A1	D1	Yes	A2	D1	Yes	C1	D3	No	D3	D2	Yes	B4	C4	No
A2	D2	Yes	B2	A1	Yes	C1	C4	No	D3	A1	Yes	C3	C4	No
C1	A1	Yes	B2	B1	Yes	C1	D4	No	D3	A2	Yes	C3	D4	No
D1	B1	Yes	B2	C1	Yes	C1	A4	No	D3	D1	Yes	C3	B3	No
C2	A2	Yes	B2	D1	Yes	C2	C3	No	D3	D2	Yes	C3	B3	No
D2	B2	Yes	C2	A1	Yes	C2	D3	No				C3	A4	No
D1	A1	Yes	C2	B1	Yes	C2	C4	No				C4	C3	No
D2	A2	Yes	C2	C1	Yes	C2	D4	No				C4	D3	No
			C2	D1	Yes	C2	A4	No				C4	A4	No
			D2	A1	Yes	D1	D3	No				C4	B4	No

D2	B1	Yes	D1	A3	No	C4	D4	No
D2	C1	Yes	D1	D4	No	D3	D4	No
D2	D1	Yes	D1	A4	No	D3	A4	No
			D1	B4	No	D3	C3	No
			D2	D3	No	D3	C3	No
			D2	A3	No	D3	B4	No
			D2	D4	No	D4	D3	No
			D2	A4	No	D4	A3	No
			D2	B4	No	D4	B4	No
						D4	C4	No
						D4	A4	No

Phase 3: Equivalence Testing. A standard Match-to-Sample procedure was used to test baseline relations and derived relations (i.e. symmetry, equivalence, see Table 26 for a summary) in a randomized block with a total of 96 test trials.

Table 26: Trial types per relation type were presented in testing

Directly Trained				
	A1B1	B1C1	C1D1	
	A2B2	B2C2	C2D2	
Symmetry				
	B1A1	C1B1	D1C1	
	B2A2	C2B2	D2C2	
1 Node				
	A1C1	B1D1	C1A1	D1B1
	A2C2	B2D2	C2A2	D2B2
2 Node				
	A1D1	D1A1		
	A2D2	D2A2		

The on-screen instruction was identical to the training phase with additional paragraph:

> "This is a test phase. Therefore, the computer will NOT tell you whether you have made the correct or wrong choice, but it is possible to get every trial correct based on what you have previously learned. Remember, your task is to match the foreign word on the bottom that goes with the one at the top. Please try to get as many correct as possible. Please report to the experimenter when the computer asks you to do so. Thank you, and good luck."

70% above accuracy was employed as mastery criterion in equivalence test phase to accommodate a low pass rate in Condition 8 (50% Reinforcement Equal no. of presentations Simultaneous Acquisition).

4.2.2 Results

102 participants reached more than 70% accuracy during the equivalence test phase, 147 participants failed to reach the mastery criterion of 70% accuracy, of which 11 participants made more than 30% errors on the lexical decision task. Table 27 depicts a generally low pass rate across conditions, with the exception of Condition 6 (50% Reinforcement Unequal no. of presentations, Simultaneous Acquisition), while Condition 8 (50% Reinforcement Equal no. of presentations, Simultaneous Acquisition) had the lowest pass rate with merely above 30% accuracy, despite relatively larger sample size. This might suggest a role of no. of stimulus presentations in equivalence formation when reinforcement and time of acquisition is kept constant. The unbalanced participant numbers across conditions did not affect the ratio of success in equivalence test performance in the current study.

Table 27: Number of passed and failed participants and pass ratio in each condition.

Conditions	No. of Passed	No. of Failed	Pass Rate
1(100%ReinfUneqPresentPhasedA)	16	17	48%
2(100%ReinfUneqPresentSimultA)	15	20	43%
3(100%ReinfEquPresentPhasedA)	11	17	39%
4(100%ReinfEquPresentSimultA)	14	23	38%
5(50%ReinfUneqPresentPhasedA)	8	13	38%
6(50%ReinfUneqPresentSimultA)	10	4	71%
7(50%ReinfEquPresentPhasedA)	8	8	50%
8(50%ReinfEquPresentSimultA)	20	45	31%

4.2.2.1 Passed Participants

Priming effects are typically calculated by subtracting reaction times (RTs) or percentage errors for target stimuli that follow related primes from RTs or percentage errors for targets that follow either unrelated primes or neutral primes. Because the error data almost always yields either no effects or the same priming effects as the RT data (Neely, 1991), the current analyses will only focus on RTs. RTs obtained from correct responses in the lexical decision task that had less than 30% errors, were grouped according to trial types (i.e., CrossClass, and Equivalence Class 1) in each condition, any value that was lower than 199 milliseconds (ms.), or exceeded 1000 ms. were considered outliers, and therefore, excluded from the following analyses.

RTs were faster for related stimulus pairs than pairs from different equivalence classes across Conditions 1 to 4, and 7 (100%ReinfUneqPresentPhasedA, 100%ReinfUneqPresentSimultA, 100%ReinfEquPresentPhasedA, 100%ReinfEquPresentSimultA, 50%ReinfEquPresentPhasedA). Interestingly, with the exception of Condition 7, where only time of acquisition was optimal, the rest of the conditions that showed faster RTs for related stimulus pairs all shared a common feature, that is, their reinforcement probability were kept relatively high, regardless of the other two factors (i.e. no. of presentations and time of acquisition). See Figure 19 for more details.

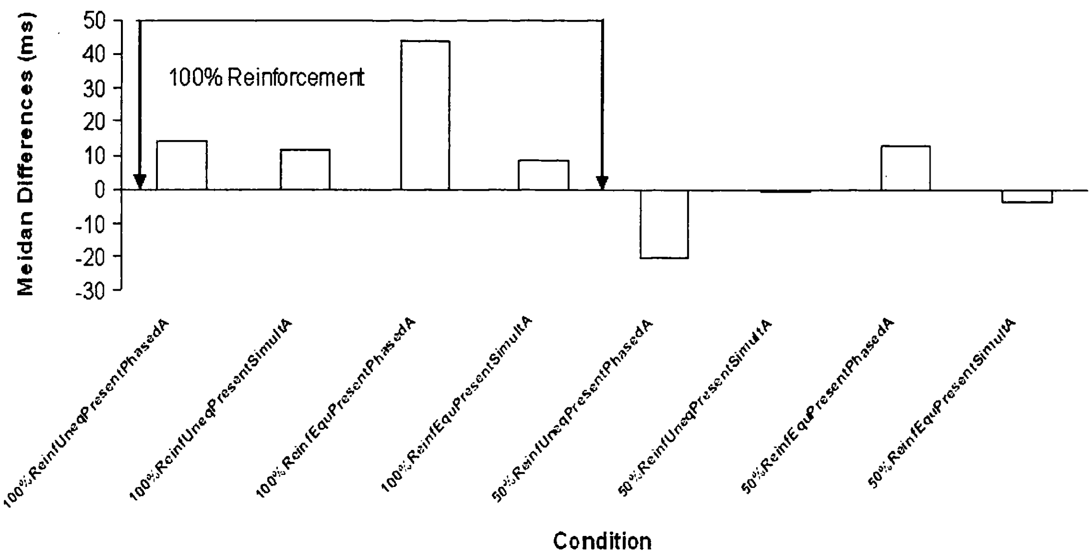


Figure 19: Group medians across the 8 conditions for the participants who passed the equivalence test.

A mixed 2 x 8 ANOVA with condition (100%ReinfUneqPresentPhasedA, 100%ReinfUneqPresentSimultA, 100%ReinfEquPresentPhasedA, 100%ReinfEquPresentSimultA, 50%ReinfUneqPresentPhasedA, 50%ReinfUneqPresentSimultA, 50%ReinfEquPresentPhasedA, 50%ReinfEquPresentSimultA) as between-subject factor, trial types (Cross Class vs. Class1) as within-subject factor was conducted in order to examine medians across conditions, trial types and any possible emergent interactions. Significant main effects were found, between RTs from related pairs and pairs from different equivalence classes, $F(1,94) = 5.783$, $p= .018$, and for condition, $F(7,94) = 2.336$, $p= .03$. A significant interaction between trial type and condition also emerged, $F(7, 94) = 3.103$, $p=.005$. However, post-hoc tests subsequently indicated that no significant differences were found between conditions, except between Condition 3 (100%ReinfEquPresentPhasedA) and Condition 7 (50%ReinfEquPresentPhasedA) ($p=.011$). Pairwise comparisons for trial type and condition interaction also showed no significance between trial types within condition with the exception of Condition 3 (100%ReinfEquPresentPhasedA), ($p<.001$), indicating that responses to equivalence related pairs were significantly faster than to nonequivalence related pairs in Condition 3.

4.2.2.2 Failed Participants

RTs from participants who failed to reach 70% or above accuracy in the equivalence test revealed a very different pattern. Priming effects were observed in Condition 3, 4 and 5, (100%ReinfEquPresentPhasedA, 100%ReinfEquPresentSimultA, 50%ReinfUneqPresentPhasedA) but subsequent statistical analysis indicated that these differences did not reach significance. See Figure 20 for more details.

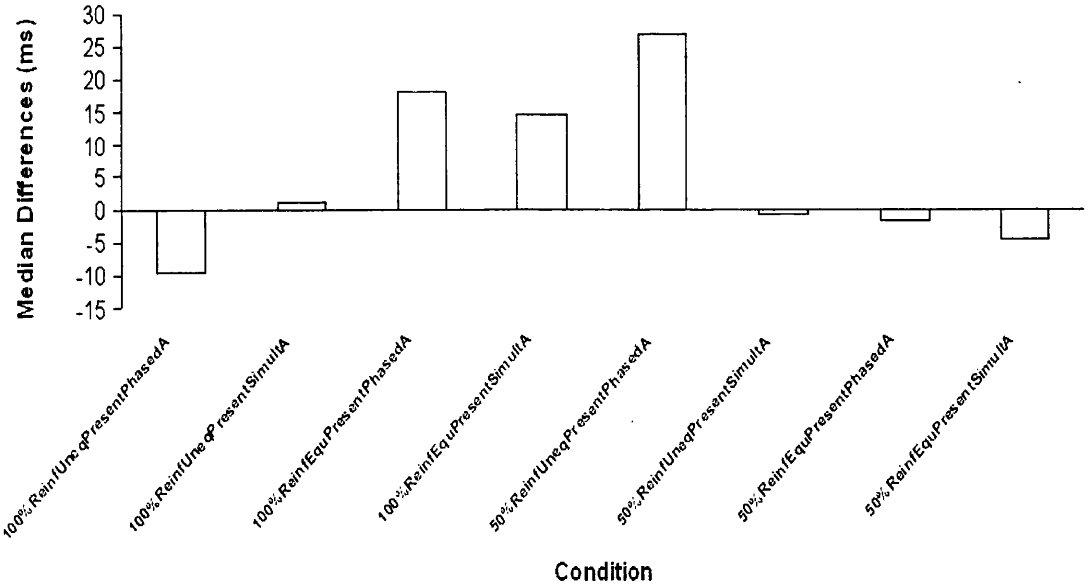


Figure 20: Group medians across 8 conditions in failed participants

In summary, less than half of the participants (102/249) reached above the 70% mastery criterion, where Condition 6 (50%ReinfUneqPresentSimultA) had the highest yield, whilst Condition 8 (50%ReinfEquPresentSimultA) had the lowest yield. For passed participants, subtracting RTs from CrossClass to Class 1 demonstrated overall priming across Conditions 1 to 4, and 7 (100%ReinfUneqPresentPhasedA, 100%ReinfUneqPresentSimultA, 100%ReinfEquPresentPhasedA, 100%ReinfEquPresentSimultA, 50%ReinfEquPresentPhasedA) with optimal time of acquisition in Condition 7 (50%ReinfEquPresentPhasedA) and high reinforcement from Condition 1 to 4 (100%ReinfUneqPresentPhasedA, 100%ReinfUneqPresentSimultA, 100%ReinfEquPresentPhasedA, 100%ReinfEquPresentSimultA), regardless of manipulation of no. of stimulus presentations and time of acquisition. Follow-up tests showed no differences amongst conditions with the exception of the difference between Condition 3 (100%ReinfEquPresentPhasedA) and 7 (50%ReinfEquPresentPhasedA), indicating reinforcement is the most salient factor that influences the strength of priming, when time of acquisition and no. of stimulus presentations were held constant. No interaction between trial type and condition was found, with the exception of the RTs between equivalently related pairs and non equivalently related pairs in Condition 3 (100%ReinfEquPresentPhasedA), indicating 100% reinforcement phased acquisition, but equal no. of stimulus presentations is capable of establishing the strongest connections between nodes, when compared to other conditions. For the failed participants, median RT differences showed priming effects from Condition 3 to 5 (100%ReinfEquPresentPhasedA, 100%ReinfEquPresentSimultA, 50%ReinfUneqPresentPhasedA), but these differences were not strong enough to reach statistical significance, suggesting that equivalence formation is a prerequisite for priming to occur.

4.2.3 Discussion

The current experiment supported the previous findings of Barnes-Holmes, et al.,'s 2005 Exp2, in that participants RTs were faster for related stimulus pairs than pairs that were from different equivalence classes, only in cases where participants formed equivalence at the end of the experiment. The differences found between passed and failed participants again supported the argument that derived relations, rather than directly reinforced stimulus relations alone, provide a behavioural model of semantic networks (Barnes-Holmes, et al 2005; Barsalou, 1999; Deacon, 1997; Hayes & Bisset, 1998).

All priming effects that were observed emerged during conditions where reinforcement levels were kept relatively high (100%), regardless of the other two factors. These findings suggest that reinforcement alone as studied in the behavioural literature is sufficient to produce derived stimulus relations that resemble what cognitive researchers refer to as semantic network growth (Imam 2001, 2003, and 2006). Additional support for this postulate emerged from the analysis of the cross condition interactions, that is, the priming effect was significantly larger when reinforcement was kept high, even when the other two factors were kept constant. Thus, indicating, that reinforcement is the most important single factor in predicting the connective strength between a newly introduced word and the existing network, which might mask any effects of when and how frequently the word was encountered. These findings coincide with Sidman's reinforcement contingency theory, that is, contingencies of reinforcement are the only major contributing factor imposed on a simple discrimination that results in between-class equivalence performance. Thus, adding further support to the assumption that equivalence (Sidman, 1994) or derived relations (Hayes, et al, 2001) might be usefully viewed as the primary component unit in a semantic network.

The observed priming effect in Condition 7 (50%ReinfEquPresentPhasedA), but not in Condition 8 (50%ReinfEquPresentSimultA), also suggested that words learned earlier were more strongly connected than words learned simultaneously when reinforcement and no. of stimulus presentations are held constant. Moreover, across 8 conditions, Unequal no. of stimulus presentations yielded the highest pass rate (71%); while stimulus pairs that were 50% in reinforcement and equal in no. of presentations and simultaneous acquisition produced the lowest yield (31%), despite the largest sample size (65) amongst conditions. This might suggest that the no. of presentations is not strong enough to express itself in a lexical decision task, even when reinforcement and time of acquisition were controlled for, but can be expressed subtly in a test for derived stimulus relations.

Although statistical significance was observed on all main comparisons and their interactions, follow-up tests demonstrated far less significance between specific variables than predicted. This lack of significance in follow-up tests, and in the 100% reinforcement, unequal no. of stimulus presentations, phased acquisition condition, might be due to low power in the current study. First, low numbers of participants formed equivalence in every condition and more than half of them failed to form equivalence after the lexical decision task. Second, the low (70%+) accuracy criterion on the equivalence test limited the findings (i.e. participants who scored below 85% were considered to have failed equivalence in Chapter 3).

Despite the manipulation of reinforcement, no. of stimulus presentations and time of acquisition, the low pass rate in the current study may also be a result of procedural differences between the current study and that employed by Barnes-Holmes's Exp2. The current experiment applied a fixed number of training (180) trials during the equivalence training phase, unlike in Barnes-Holmes's Exp2, in which 24 consecutively correct trials were required before completing one training exposure out of 10. As a result of this intense training, their participants were exposed to more baseline relations (413 – 424 trials) than participants in the current study, thus, their baseline relations were more firmly established prior to the derived relations test.

There were other differences between the two studies. There was a longer SOA (600ms) in the lexical decision task in the current study, with the same 200ms latency of the prime stimulus as in Barnes-Holmes's Exp2, the blank screen between the prime and the target was increased from 50ms to 400ms in the current study. Despite the different SOAs used, in the cognitive literature both SOAs would be considered as short (Neely, 1991). Moreover, eight instead of six novel nonsense words were used in the current lexical decision task, with balanced trial types in the presentation of stimulus pairs from each different equivalence class, which resulted in an increased overall number of trials being presented. Finally, each trial type was presented only once in the current lexical decision task as opposed to twice in Barnes-Holmes's Exp2, which might affect the baseline relation performance during testing.

In conclusion, this is the first empirical study to investigate the interaction between reinforcement, no. of stimulus presentations and time of acquisition in semantic network growth. The findings from the current experiment were generally consistent with the semantic network literature, and the current work adds to the burgeoning behavioural literature that implicates reinforcement as the most salient single factor influencing equivalence class formation/the development of semantic networks. Thus, the current work adds to the emergent literature that bridges the gap between semantic network research and stimulus equivalence (Barnes-Holmes, et al, 2005; Haimson et al., 2009; Yorio, et al., 2008).

Chapter 5: Neurological Effects of Reinforcement and Nodal Number in Equivalence Class Formation

5.1 General Introduction

The findings from Chapter 4 indicated that reinforcement probability played a more significant role in determining the connection strength of trained relations in an equivalence class than time of acquisition or no. of stimulus presentations. In particular, participants' RTs were significantly different between 100% versus 50% reinforcement when holding time of acquisition and no. of stimulus presentations constant. The current chapter sought to extend this finding by introducing EEG as an additional dependent measure (see Section 1.4). The aim of the addition measure was to directly test the neurological processes underlying semantic network growth.

Recently, the possible overlap between behavioural and cognitive explanations of semantic processing (see Chapter 4) has opened a new avenue for integrating neurological techniques into behaviour analysis. Barnes-Holmes et al. (2005) demonstrated that mediated priming emerges in equivalence related pairs, but not in non equivalence related pairs (Exp's 1 + 2). In Experiment 3, EEG was recorded in order to determine whether the N400 commonly observed for a semantic mismatch would also differentiate between non-equivalence and equivalence stimulus relations on a lexical decision task. The procedure was similar to that employed by Barnes-Holmes et al, (2005: Expt 2). In Experiment 3, two 4-member equivalence classes were trained using a MTS procedure with trials introduced quasi-randomly. After repeated MTS training (10 times), a two-word lexical decision task similar to that used by Hayes and Bisset (1998) with a 100ms SOA was used, EEG and behavioural data were recorded simultaneously during the lexical decision task. Participants RTs were consistent with Experiment 2, that is, participants responded significantly faster to equivalence than non-equivalence stimulus pairs, and faster to directly trained than derived stimulus pairs. The averaged EEG signals across the 20 participants who formed equivalence indicated significantly more negativity for the N400 in non equivalence relations compared to derived relations on all electrode sites monitored on the left hemisphere (C3, P3, T3, T5, & O1) and significantly more negativity of the N400 in equivalence pairs compared to directly trained pairs in 4/5 electrode sites with the exception of the site C3. There was some evidence of differences between directly trained pairs and both the equivalence and non equivalence pairs on the right hemisphere, but this was not as robust as for the left hemisphere. This experiment provided the first empirical evidence that the N400 commonly associated with semantic processing can be used to differentiate equivalence and non-equivalence relations.

Further evidence of reflexivity was provided by Yorio and his colleagues (2008). They trained participants in two 3-member equivalence classes (AB, AC) on a delayed MTS procedure using figures of artificial objects. Participants were tested for reflexivity, equivalence and non-equivalence relations using a one-word lexical decision paradigm with 2500 ms SOA. Event Related Potentials (ERPs) not only demonstrated the greater negativity of the N400 on non-equivalence compared to equivalence relations, but also indicated a greater negativity in equivalence relations than on reflexivity relations (e.g. A = A) on a timeframe of 150 – 250ms. These researchers also found a larger positivity around 300 milliseconds after stimulus onsets in equivalence rather than non-equivalence relations, which might resemble the P300 commonly reported in categorization studies (Azizian, Freitas, Watson, & Squires, 2006). However, Tabullo, Yorio, Leguizamon, and Segura (2008), replicated the emergence of the N400 in a category learning task, and suggested that the N400 could be a neural marker of categorization and decision making.

Evidence of successful equivalence formation results not only from simple discrimination training, but also from emergent relations in the test themselves (Sidman, Kirk, & Willson-Morris, 1985) was demonstrated by Haimson and his colleagues (2009). In the Experiment 1, the N400 was demonstrated using English word-pairs with 1900 ms SOA. In the Experiment 2, three 6-member equivalence classes were trained using a fixed-sample MTS procedure with a serial training structure (AB introduced first, then AC, then AD, then AE, then AF), with a gradual increase in the number of comparisons in a single trial (i.e., A1 only had one comparison, B1, then the number of comparisons was increased to B1 and B2, until all three comparisons were presented with A1). The A stimuli were trigrams, while the remainder of the stimuli were figures of artificial objects. Participants were tested for emergent relations (symmetry and transitivity) across two different groups (i.e., EEG tested either before or after the test). The findings indicated a robust N400 for the participants who received EEG testing after the test of emergent relations, but not prior to the test. Moreover, they argued that a gradual emergence of the N400 in repeated testing was an analogue for delayed emergence of equivalence commonly observed in behavioural data (this issue will be revisited in Experiment 5.2). However, the authors only focused on the N400, and did not report on the P300-like waveforms in Figure 1 (p.248) that could reflect the outcome of categorization after equivalence formation.

The general aim of Experiments 7 and 8 was to provide preliminary neurological evidence of a behavioural account of semantic network growth. As demonstrated in Experiment 6, reinforcement is one important contributor in determining the strength of connections between nodes (i.e., relatedness among stimuli). Experiment 1- 4 also showed that the number of trial presentations affected the success of equivalence performance. Experiment 7 sought to investigate the interactions of reinforcement and number of trial presentations at both the behavioural and neurological level. To that end, Experiment 7 employed a 2x2 between-subject design, with levels of reinforcement (high vs. equal), and levels of trial presentation (high vs. low) as independent variables. It is predicted that, if an equivalence relation is indeed the basic structure of semantic relations, a greater negativity of the N400 typically reported in semantic mismatching would also be found in non equivalence related stimulus pairs compared to equivalence related pairs. Specifically, high reinforcement, high trial presentation will lead to the strongest connections between nodes, therefore, the fastest RTs and the least negativity (on the N400) for equivalence relations, whereas, equal reinforcement and low trial presentation will lead to the weakest connections between nodes, the longest RTs and the largest negativity for equivalence relations. Ideally, a greater positivity on the P300 will be found in equivalence related pairs compared to non equivalence related pairs if the P300 reflects categorization post-equivalence formation (Experiment 7). Experiment 8 employed a within-subject design with trial type as the independent variable. If the stimulus control based on node number within a given equivalence class is a contributing factor, one would expect to found a greater P300 in stimulus pairs with more nodes than less nodes in between if P300 reflects the degree of source allocation. Greater negativity of N400 in 4-node than 1-node relation could confirm the differential control exerted by node numbers amongst equivalent stimulus.

5.2 Experiment 7

5.2.1 Method

5.2.1.1 Participants

58 healthy right-handed university students with normal or corrected to normal vision, participated in the current experiment (18 male; 40 female), ranging in age from 18 to 36 years (m = 22.15, SD = 4.07). Most participants were students from Swansea University and Swansea Metropolitan University. Psychology undergraduates were either given subject pool credits at the end of the experiment, or one £10 payment for their participation. All of the participants were naïve about the purpose of the experiment, and were assured that they could withdraw without penalty at any time during the experiment. The study was approved by the Department of Psychology, Swansea University Ethics Committee.

The exclusion of left handed participants or those with any drug and alcohol dependency was to control for variables (e.g. abnormality of brain activity in long term drug and alcohol abuse) that were beyond the interest of the current research.

5.2.1.2 Apparatus and Materials

The experiment was conducted in a sound-attenuated, electrically shielded cubicle in the electrophysiology laboratory in the Department of Psychology at Swansea University. The apparatus and stimuli used were identical to those in Experiment 6. In order to record the EEG measures during the lexical decision task, an Active Two Mark II system with control software (ActiView™ BioSemi) and an electrode cap (32-channel) were employed. The Active Two Mark II system and ActiView™ BioSemi were controlled by a Dell desk top computer with a Pentium 4 processor. All of the hardware and software described above were manufactured and supplied BioSemi, B. V., 1054SC Amsterdam, Netherlands. Finally, the ERP data were analyzed using analysis software (BESA research Version 5.3), which was supplied by MEGIS Software GmbH, Gräfelfing, Germany. Materials were identical to those used in Experiment 6.

5.2.1.3 Procedure

Before starting the experiment, each participant was given an information sheet which outlined certain medical conditions that would involve exclusion from participating in the current study. Edinburgh Handedness Inventory (Oldfield, 1971, see Appendix 1) was employed to assess participants' handedness with questions such as "which hand do you prefer to use when writing?" The procedure was similar to that employed in Experiment 6, with the exceptions that: only condition 1 (HighR) and 5(EqualR) were employed in order to evaluate the effect of reinforcement, while holding time of acquisition and no. of stimulus presentaions constant. 40 Participants were randomly assigned to the two experimental conditions (HighR and EqualR) with repeated training (i.e., two exposures) to stabilize baseline performance. The remaining 18 participants were trained only once, due to subject pool credit and time restrictions, hence, were identical to Condition 1 and 5 in Experiment 6. The study was a 2 (High [100%] vs. Equal [50%] reinforcement) x 2 (High [twice] vs. Low [once] trial presentation) between subjects design. For ease of communication the trial types will be denoted as follows: HRLowTrial, HRHighTrial, ERLowTrial, ERHighTrial.

5.2.1.4 EEG Recording

EEG was applied as an additional measure to the behavioural response time measure during the lexical decision task. Voltage recordings were performed on the scalp in accordance with the 32+2 system in Fp1-2, AF3-4, F3-4-Z, F7-8, FC1-2, FC5-6, C3-4-Z, CP1-2, CP5-6, T7-8, P3-4-Z, P7-8, PO3-4, O1-2-Z, plus CMS and DRL as reference channels from a 32+2 channel elastic Electro-cap. The bandwidth was set between 0.3 and 40 Hz with a sampling rate of 16384 Hz. All electrode impedances were at or below 50 kΩ. The EEG was continuously collected and edited off-line with BESA (Version: research 5.3). Epochs of 900ms with a pre-target stimulus time of 200 ms were averaged. Horizontal and vertical electro-oculogram (EOG) were automatically corrected with other artefacts by BESA. All data were average referenced. Fourteen sites were further analyzed statistically (F3-4, FC1-2, C3-4, CP1-2, P3-4-Z, O1-2-Z). These sites were chosen because most of they (i.e. F3-F4, C3-4, P3-4, O1-O2), were found closely associated with N400 (Barnes-Holmes et al, 2005, Yorio, et al, 2008, Haimson, et al, 2009), less noisy as FZ and CZ. Averaged ERPs to primes and targets that were directly trained (e.g., A1-B1), related through equivalence class (e.g. B1-A1, C1-A1), or unrelated through class inconsistency (e.g. CrossClass trials, A1-A2), or unrelated through equivalence and previously unseen stimuli (e.g. Class-nonClass, A1-A3) were obtained from correct response trials for each participant, above 85% trials in each condition.

5.2.2 Results

Anyone who reached the 80% or above accuracy mastery criterion in the equivalence test was treated as having demonstrated equivalence performance. 17 participants completed the HRHighTrial condition (12 passers), 23 the ERHighTrial condition (13 passers), 10 the HRLowTrial condition (8 passers) and 8 the ERLowTrial condition (6 passers). All conditions reached 70% above pass rate, except ERHighTrial condition with a pass rate at slightly above chance level.

5.2.2.1 Passed Participants

Priming effects are typically calculated by subtracting reaction times (RTs) or percentage incorrect responses for target stimuli that follow related primes from RTs or percentage incorrect responses for targets that follow either unrelated primes or neutral primes. Because the accuracy on incorrect responses almost always yields either no effects or the same priming effects as the RT data (Neely, 1991), the current analyses will only focus on RTs. RTs obtained from correct responses in the lexical decision task that had less than 30% incorrect responses, were grouped according to trial types (i.e., CrossClass, and Equivalence Class 1) in each condition, any value that was lower than 199 milliseconds (ms.), or exceeded 1000 ms. were considered outliers (averaged percentages ranged from 3% to 5.4%), and therefore, excluded from the following analyses.

RTs were faster for related stimulus pairs than pairs from different equivalence classes across all conditions, except for the ERLowTrial condition. Amongst the three conditions that showed priming effects, there was interesting interactions between reinforcement and trial presentation, that is, greater priming in the HRHighTrial condition, followed by the ERHighTrial condition, and lower priming in the HRLowTrial condition. (see Figure 21). However, non parametric statistics (i.e., Wilcoxon, Mann-Whitney) showed none of those differences to be significant ($p>0.05$).

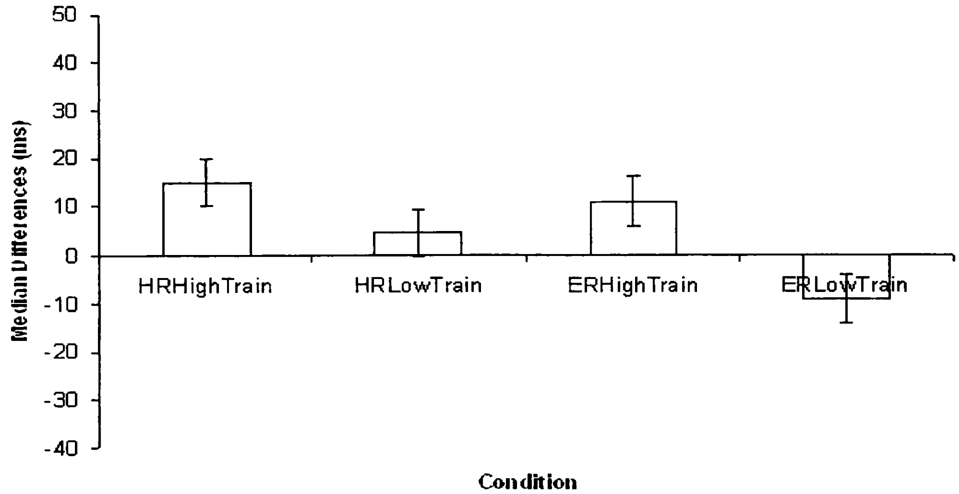


Figure 21: Median of mean differences and 95% confidence intervals across all conditions for the passed participants

The averaged waveforms across all conditions indicated that signals of the directly trained trials were the most affected by experimental manipulations across trial types with relatively more deviations, this might be due to the fact that fewer trials were available to be averaged (i.e., only 3 trials in each participant); the waveforms generated by class, non-class trials (i.e.,. A1A3) was almost identical to the waveform generated by class inconsistent trials (i.e., CrossClass, A1A2), therefore, only CrossClass trials served as a comparison to directly trained and equivalence related trials in statistical analysis. It was also found a large positivity around 300 ms after target onsets in the frontal lobe area; therefore, statistical analysis was performed across these two timeframes: 250 – 350 ms, and 350 – 550 ms.

5.2.2.1.1 250 – 350 milliseconds

For the 250 – 350 ms timeframe, three mixed ANOVAs with trial type (directly trained, equivalent, and class inconsistent), electrode sites (F3-4, FC1-2, C3-4) as repeated factors and condition (HRHighTrial, HRLowTrial, ERHighTrial, ERLowTrial) as the between-subject factor were calculated across 3 frequency ranges: Theta (4-8 Hz), Alpha (8-13 Hz), and Beta (13-22 Hz), respectively. A significant main effect was found from 4 to 8 Hz amongst sites, $F(5, 145) = 14.741$, $p < .0001$. The main effect for trial types was not significance ($p = .094$). A significant main effect was found from 8 to 13 Hz amongst conditions, $F(3, 29) = 2.935$, $p = .05$;

electrode sites, F (5, 145) = 11.877, p<.0001. And there was a significant interaction between trial type and condition, F (6, 58) = 2.889, p=.016. Subsequent post-hoc tests across conditions did not find any significant differences. Significant results obtained in follow up tests between the trial type and condition interaction are shown in Table 28. No significant differences were found for the Beta frequency (13-22Hz).

Table 28: Results of Bonferroni pairwise tests comparing HRHighTrial, ERHighTrial, HRLowTrial, ERLowTrial across direct trained (DT), equivalent (EQ), and class inconsistent (CC) trials from 250 to 350 ms in 8 to 13 Hz.

Comparisons	Significant Results
HRHighTrial vs. HRLowTrial	DT (p=.013)
HRLowTrial vs. ERLowTrial	DT (p=.004)
HRHighTrial vs. HRLowTrial	CC (p=.032)
DT vs. CC	HRLowTrial (p=.018), ERLowTrial (p=.05)

5.2.2.1.2 350 – 550 milliseconds

For the 350 to 550 ms timeframe, three mixed ANOVAs with trial type (directly trained, equivalent, and class inconsistent), electrode sites (F3-4, FC1-2, C3-4, CP1-2, P3-4-Z, O1-2-Z) as repeated factors, condition (HRHighTrain, HRLowTrain, ERHighTrian, ERLowTrain) as between-subject factor were calculated in 3 frequency range: Theta (4-8 Hz), Alpha (8-13 Hz), and Beta (13-22 Hz), respectively. Significant main effects were found between 4 to 8 Hz in trial type, F (2, 56) =5.134, p=.009; sites, F (13, 364) = 5.313, p<.0001. Subsequent pairwise comparisons amongst trial types demonstrated significant differences between direct trained and class inconsistent trials (p=.029), between equivalent and class inconsistent trials (p=.045). Significant main effects were also found between 8 to 13 Hz amongst electrode sites, F (13, 364) =17.839, p<.0001; Significant interaction was found amongst trial type, sites and condition, F (78, 728) =1.385, p=.02; Results of the bonferroni pairwise tests on the 3-way interaction was presented in Table 29. Approaching significance was also found in interaction between sites and condition (p=.082).

Table 29: Results of Bonferroni pairwise tests comparing HRHighTrial (1), ERHighTrial (2), HRLowTrial (3), ERLowTrial (4) across direct trained (DT), equivalent (EQ), and class inconsistent (CC) trials from 350 to 550 ms in 8 to 13 Hz.

	Statistical significance								
	Left sites				Middle	Right sites			
Pairwise	FC1	CP1	P3	O1	OZ	O2	CP2	C4	FC2
1 vs. 2		EQ[b]				DT[b]	DT[a]		
1 vs. 3		EQ[a]			DT[b]	DT[b]			
2 vs. 4				DT[a]	DT[b]		DT[b]	DT[b]	DT[b]
3 vs. 4			CC[a]	DT[a]	DT[a]			DT[b]	DT[b]
DT vs. CC						2[b]3[b]			
EQ vs. CC	2[a]	3[b]							

[a] p<.05.

[b] Approaching significance (p<.097)

In summary, participants who formed equivalence demonstrated priming effects across the HRHighTrial, ERHighTrial and HRLowTrial conditions, with the HRHighTrial condition demonstrating the largest priming effect and the HRLowTrial condition demonstrating the weakest priming effect. However, no statistical significance was found either within or between conditions. Further analysis of mean amplitude in two timeframes in both Theta and Alpha frequencies showed interesting and complex results across conditions. In 250 to 350 ms timeframe trial types in Theta frequency were approaching significance. Within Alpha frequency range, significant main effect was found amongst conditions, but no significance was found between conditions in post-hoc tests. Significant interactions were found between trial type and conditions. Bonferroni pairwise tests indicated significant differences between the HRHighTrial and HRLowTrial conditions and between the HRLowTrial and ERLowTrial conditions for directly trained trials. A significant difference was also found between the HRHighTrial and HRLowTrail conditions for class inconsistent trials. Differences were significant between directly trained and class inconsistent trials for both the HRLowTrial and ERLowTrial conditions.

For the 350 to 550 timeframe, significant differences of mean amplitudes were found between directly trained and class inconsistent trials and between equivalence and class inconsistent trials, regardless of experimental manipulations in the Theta frequency. Within Alpha frequency range, mean amplitudes in directly trained trials were significantly different between the HRLowTrial and ERLowTrial conditions on sites O1, and OZ, approaching significance at sites C4 and FC2. These differences were also found between the ERHighTrial and ERLowTrial conditions at sites O1, OZ, CP2, C4, and FC2, between HRHighTrial and ERHighTrial conditions on sites O2 and CP2, between HRHighTrial and HRLowTrial condition on sites OZ and O2, however the majority of the sites were only approaching significance. Moreover, mean amplitudes were significantly different between equivalent and class inconsistent trials at site FC1 in ERHighTrial condition only (see Figure 22), approaching significance at site CP1 in the HRLowTrial condition. Differences between directly trained and class inconsistent trials were approaching significance at site O2 on both the ERHighTrial and HRLowTrial conditions. Right hemisphere sensitivity to the effects of the experimental manipulation on directly trained relations was also observed.

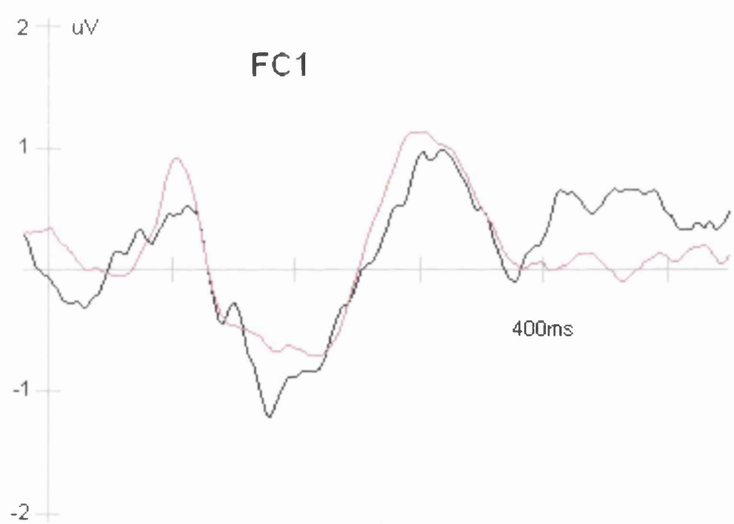


Figure 22: Grand average waveforms across all passed participants in site FC1: the black line indicates equivalent related trials, the red line indicates class inconsistent trials.

5.2.2.2 Failed Participants

5 participants failed in HRHighTrial condition, 10 failed in ERHighTrial, in which 2 had more than 30% errors in lexical decision task. HRLowTrial and ERLowTrial each had 2 failed participants. RTs obtained from correct responses in the lexical decision task that had less than 30% errors, were grouped according to trial types (i.e., CrossClass, and Equivalence Class 1) in each condition, any value that was lower than 199 milliseconds (ms.), or exceeded 1000 ms. were considered outliers, and therefore, excluded from the following analyses. RTs were faster for equivalence trials than class inconsistent trials in the HRLowTrial condition only (it is important to note that this was calculated from the only 2 participants). Figure 23 depicted RTs and 95% confidence intervals across all conditions for the failed participants.

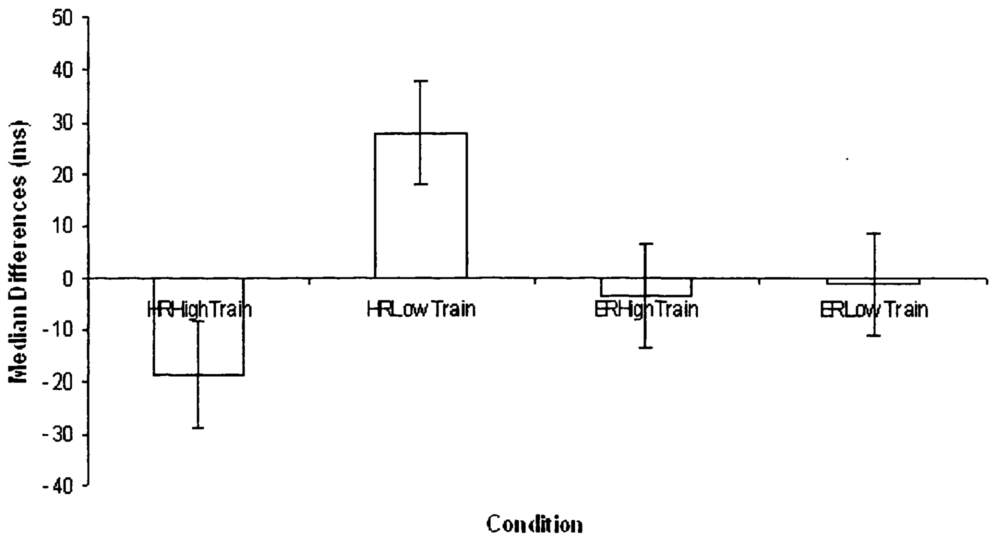


Figure 23: Median of mean differences and 95% confidence intervals across all conditions for failed participants

5.2.3 Discussion

Analysis of RTs from the current experiment indicated that the strength of priming is sensitive to the impact of reinforcement and trial presentation. The low reinforcement low trial presentation condition did not show a priming effect for the passed participants. This finding provided the first empirical evidence for the assumption that reinforcement and number of trial presentations are contributing factors in the formation of semantic networks. The lack of mediated priming in failed participants replicated the finding in Chapter 4 and are consistent with Barnes-Holmes et al., (2005)'s Exp2, confirming that derived stimulus relations resembles what cognitive psychologists refer to as "mediated priming".

Analysis of mean amplitude for the passed participants, demonstrated a significant difference between class inconsistent trials and stimulus equivalence trials in the 4 – 8 Hz frequency range regardless of experimental manipulation. However, no N400 effect was reported in the 8 – 13 Hz frequency range, except site FC1. The lack of robust N400 effects in the 8 – 13 Hz suggested that the current procedure might have masked the robust effect of the N400. For example, all experiments that reported the N400 in the equivalence literature, employed a simple experimental design without manipulating reinforcement and trial presentation as in the current study. Nevertheless, the analysis of mean amplitude provided positive results suggesting that direct priming was the most sensitive when comparing equivalence and class inconsistent relations, with a greater negativity in high rather than low reinforcement, low number of trial presentations conditions 400 ms after the target onset on sites O1 and OZ. The N400 is normally associated with a semantic mismatch, here it seems to reflect the level of reinforcement imposed on stimulus relations. Additionally, the mean amplitudes of directly trained trials were significantly different between high trial presentation and low trial presentation in the high reinforcement conditions, between high rather than low reinforcement in low trial presentation conditions 300 ms after target onset. These differences might suggest that the expression of the effect of trial presentation on direct priming requires a high level of reinforcement, whereas the expression of the effect of reinforcement requires a low number of trial presentations. Right hemisphere sensitivity to the effects of the experimental manipulations on directly trained relations were also observed, confirming that the right hemisphere is linked to learning and decision making.

5.3 Experiment 8

5.3.1 Method

5.3.1.1 Participants

42 healthy right-handed university students with normal or corrected to normal vision, participated in the current experiment (17 male; 25 female), ranging in age from 18 to 46 years (m = 21.29, SD = 4.66). Most participants were students from Swansea University and Swansea Metropolitan University, Psychology

undergraduates were given credits at the end of the experiment as fulfilling of their course requirement, and other students were offered £10 instead. All of the participants were naïve about the purpose of the experiment, and were assured that they could withdraw without penalty at any time during the experiment. The study was approved by the Department of Psychology, Swansea University Ethics Committee.

5.3.1.2 Apparatus and Materials

The Apparatus was identical to Experiment 7. Materials were identical to those employed in Experiment 1- 4.

5.3.1.3 Procedure

Unlike Experiment 7, the current experiment employed a within-subject design. Each participant was trained on two 6-member equivalence classes using a MTS procedure with a linear training structure. Feedback was presented after each trial, 20 consecutively correct trials were required for completion of AB trial type, then BC trial types were added until participants produced 20 consecutively correct, followed by the addition of CD trial types until participants produced 20 consecutively correct, followed by DE trials and so on, until 20 consecutively correct trials were produced in the last phase during which all trial types were randomly mixed in a block (i.e., completing one training cycle). The training cycle was repeated 4/5 times by each participant on the first day of the experiment and repeated another 4/5 times by each participant on the second day, resulting in 9 cycles prior to the third day of participation. Upon completion of the 10^{th} cycle of training, each participant was presented with the lexical decision task similar to Experiment 7, with the exception that each equivalence class comprises extended class members (see Table 30 for a schematic presentation of trial types presented). Participant's behavioural and EEG data were recorded simultaneously during the lexical decision task. Finally, tests of all emergent relations and baseline relations were presented randomly in a mixed block in a standard MTS format the same as in training.

Table 30: 360 trials presented during the lexical decision procedure. Pm = Prime, Tg = Target, Rp = Correct Response, A3, B3, C3, D3, E3, F3, A4, B4, C4, D4, E4, and F4 = Nonsense Word.

Within Class			Cross Class			Class-nonsense			Nonsense-class			Nonsense-nonsense		
Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp	Pm	Tg	Rp
Directly trained			A1	A2	Yes	A1	A4	No	A3	A1	Yes	A3	A4	No
A1	B1	Yes	A1	B2	Yes	A1	B4	No	A3	A2	Yes	A3	B4	No
B1	C1	Yes	A1	C2	Yes	A1	C4	No	A3	D1	Yes	A3	D3	No
C1	D1	Yes	A1	D2	Yes	A1	D3	No	A3	D2	Yes	A3	D3	No
D1	E1	Yes	A1	E2	Yes	A1	D4	No	A3	C1	Yes	A3	C4	No
E1	F1	Yes	A1	F2	Yes	A1	E3	No	A3	C2	Yes	A3	E3	No
A2	B2	Yes	B1	A2	Yes	A1	F3	No	A3	E1	Yes	A3	F4	No
B2	C2	Yes	B1	B2	Yes	A2	A3	No	A3	E2	Yes	A4	A3	No
C2	D2	Yes	B1	C2	Yes	A2	B3	No	A3	F1	Yes	A4	B3	No
D2	E2	Yes	B1	D2	Yes	A2	A4	No	A3	F2	Yes	A4	C4	No
E2	F2	Yes	B1	E2	Yes	A2	B4	No	B3	B1	Yes	A4	D4	No
Symmetry			B1	F2	Yes	A2	C4	No	B3	B2	Yes	A4	B4	No
B1	A1	Yes	C1	A2	Yes	A2	E3	No	B3	C1	Yes	A4	E4	No
C1	B1	Yes	C1	B2	Yes	A2	F4	No	B3	C2	Yes	A4	F3	No
D1	C1	Yes	C1	C2	Yes	B1	B3	No	B3	B1	Yes	B3	B4	No
E1	D1	Yes	C1	D2	Yes	B1	C3	No	B3	B2	Yes	B3	C4	No
F1	E1	Yes	C1	E2	Yes	B1	B4	No	B3	E1	Yes	B3	A3	No
B2	A2	Yes	C1	F2	Yes	B1	C4	No	B3	E2	Yes	B3	A3	No
C2	B2	Yes	D1	A2	Yes	B1	D4	No	B3	F1	Yes	B3	D4	No
D2	C2	Yes	D1	B2	Yes	B1	E3	No	B3	F2	Yes	B3	E3	No
E2	D2	Yes	D1	C2	Yes	B1	F3	No	C3	C1	Yes	B3	F4	No
F2	E2	Yes	D1	D2	Yes	B2	B3	No	C3	C2	Yes	B4	B3	No
Equivalence			D1	E2	Yes	B2	C3	No	C3	B1	Yes	B4	C3	No
A1	C1	Yes	D1	F2	Yes	B2	B4	No	C3	B2	Yes	B4	D4	No
B1	D1	Yes	E1	A2	Yes	B2	C4	No	C3	A1	Yes	B4	A4	No
C1	E1	Yes	E1	B2	Yes	B2	D4	No	C3	A2	Yes	B4	C4	No
D1	F1	Yes	E1	C2	Yes	B2	A3	No	C3	D1	Yes	B4	E3	No
A2	C2	Yes	E1	D2	Yes	B2	F4	No	C3	D2	Yes	B4	F4	No

B2	D2	Yes	E1	E2	Yes	C1	C3	No	C3	E1	Yes	C3	C4	No
C2	E2	Yes	E1	F2	Yes	C1	D3	No	C3	E2	Yes	C3	D4	No
D2	F2	Yes	F1	A2	Yes	C1	C4	No	D3	D1	Yes	C3	B3	No
A1	D1	Yes	F1	B2	Yes	C1	D4	No	D3	D2	Yes	C3	B3	No
B1	E1	Yes	F1	C2	Yes	C1	A4	No	D3	A1	Yes	C3	A4	No
C1	F1	Yes	F1	D2	Yes	C1	E3	No	D3	A2	Yes	C3	E3	No
A2	D2	Yes	F1	E2	Yes	C1	B4	No	D3	D1	Yes	C3	F4	No
B2	E2	Yes	F1	F2	Yes	C2	C3	No	D3	D2	Yes	C4	C3	No
C2	F2	Yes	A2	A1	Yes	C2	D3	No	D3	C1	Yes	C4	D3	No
A1	E1	Yes	A2	B1	Yes	C2	C4	No	D3	C2	Yes	C4	A4	No
B1	F1	Yes	A2	C1	Yes	C2	D4	No	D3	F1	Yes	C4	B4	No
A2	E2	Yes	A2	D1	Yes	C2	A4	No	D3	F2	Yes	C4	D4	No
B2	F2	Yes	A2	E1	Yes	C2	E4	No	E3	A1	Yes	C4	E3	No
A1	F1	Yes	A2	F1	Yes	C2	F3	No	E3	A2	Yes	C4	F4	No
A2	F2	Yes	B2	A1	Yes	D1	D3	No	E3	B1	Yes	D3	D4	No
C1	A1	Yes	B2	B1	Yes	D1	A3	No	E3	B2	Yes	D3	A4	No
D1	B1	Yes	B2	C1	Yes	D1	D4	No	E3	C1	Yes	D3	C3	No
E1	C1	Yes	B2	D1	Yes	D1	A4	No	E3	C2	Yes	D3	C3	No
F1	D1	Yes	B2	E1	Yes	D1	B4	No	E3	D1	Yes	D3	B4	No
C2	A2	Yes	B2	F1	Yes	D1	E4	No	E3	D2	Yes	D3	E3	No
D2	B2	Yes	C2	A1	Yes	D1	E3	No	E3	E1	Yes	D3	F4	No
E2	C2	Yes	C2	B1	Yes	D2	D3	No	E3	F2	Yes	D4	D3	No
F2	D2	Yes	C2	C1	Yes	D2	A3	No	F3	E1	Yes	D4	A3	No
D1	A1	Yes	C2	D1	Yes	D2	D4	No	F3	E2	Yes	D4	B4	No
E1	B1	Yes	C2	E1	Yes	D2	A4	No	F3	B1	Yes	D4	C4	No
F1	C1	Yes	C2	F1	Yes	D2	B4	No	F3	B2	Yes	D4	A4	No
D2	A2	Yes	D2	A1	Yes	D2	F4	No	F3	C1	Yes	D4	E3	No
E2	B2	Yes	D2	B1	Yes	D2	F3	No	F3	C2	Yes	D4	F3	No
F2	C2	Yes	D2	C1	Yes	E1	A3	No	F3	D1	Yes	E3	A3	No
E1	A1	Yes	D2	D1	Yes	E1	A4	No	F3	D2	Yes	E3	A4	No
F1	B1	Yes	D2	E1	Yes	E1	C3	No	F3	F1	Yes	E3	B3	No
E2	A2	Yes	D2	F1	Yes	E1	D4	No	F3	F2	Yes	E3	D3	No
F2	B2	Yes	E2	A1	Yes	E1	B4	No				E3	E4	No

F1	A1	Yes	E2	B1	Yes	E1	E3	No	E3	D4	No
F2	A2	Yes	E2	C1	Yes	E1	F3	No	E3	F3	No
			E2	D1	Yes	E2	B3	No	E4	F3	No
			E2	E1	Yes	E2	B4	No	E4	A4	No
			E2	F1	Yes	E2	D3	No	E4	B3	No
			F2	A1	Yes	E2	C4	No	E4	C4	No
			F2	B1	Yes	E2	A3	No	E4	C3	No
			F2	C1	Yes	E2	E4	No	E4	E3	No
			F2	D1	Yes	E2	F4	No	E4	F4	No
			F2	E1	Yes	F1	D3	No	F3	A3	No
			F2	F1	Yes	F1	D4	No	F3	D3	No
						F1	E3	No	F3	C4	No
						F1	E4	No	F3	E4	No
						F1	F3	No	F3	B3	No
						F1	F4	No	F3	C3	No
						F1	C4	No	F3	E3	No
						F2	A4	No	F4	C3	No
						F2	B4	No	F4	D4	No
						F2	C4	No	F4	D3	No
						F2	D3	No	F4	B4	No
						F2	E3	No	F4	A3	No
						F2	F3	No	F4	A3	No
						F2	F4	No	F4	E4	No

5.3.1.4 EEG recording

EEG recording was identical to Experiment 7.

5.3.2 Results

31/42 participants passed the mastery criterion of 80% in testing for emergent relations, one withdrew after the first day of training.

5.3.2.1 Passed Participants

RTs obtained from correct responses in the lexical decision task that had less than 30% errors, were grouped according to trial type (i.e., WithinClass, NoClass, DT, Symmetry, 1N, 2N, 3N, and 4N) any value that was lower than 199 milliseconds (ms.), or exceeded 1000 ms. were considered outliers, and therefore, excluded from the following analyses.

Priming effects were observed across all trial types for participants who formed equivalence. The largest priming effect was found between NoClass and directly trained trials (57.27), the smallest priming was found between NoClass and Symmetry trials (17.17) and between NoClass and 4-node trials (17.31). A series of t-tests were calculated subsequently to test the strength of priming effects, and demonstrated that the differences between NoClass and all other trial types were significant (for withinClass, $t(30) = 5.242$, $p<.0001$; for DT, $t(30) = 4.577$, $p<.0001$; for Symmetry, $t(30) = 3.812$, $p= .001$; for 1N, $t(30) = 4.475$, $p<.0001$; for 2N, $t(30) = 3.794$, $p= .001$; for 3N, $t(30) = 3.484$, $p= .002$) with the exception that there was no difference between NoClass and 4-node trials. Tests for differences between trial types did not find any significant results, however, the differences between DT and Symmetry, DT and 2N, 1N and 2N were approaching significance ($p=.071$; .078; .106, respectively) (see Figure 24)

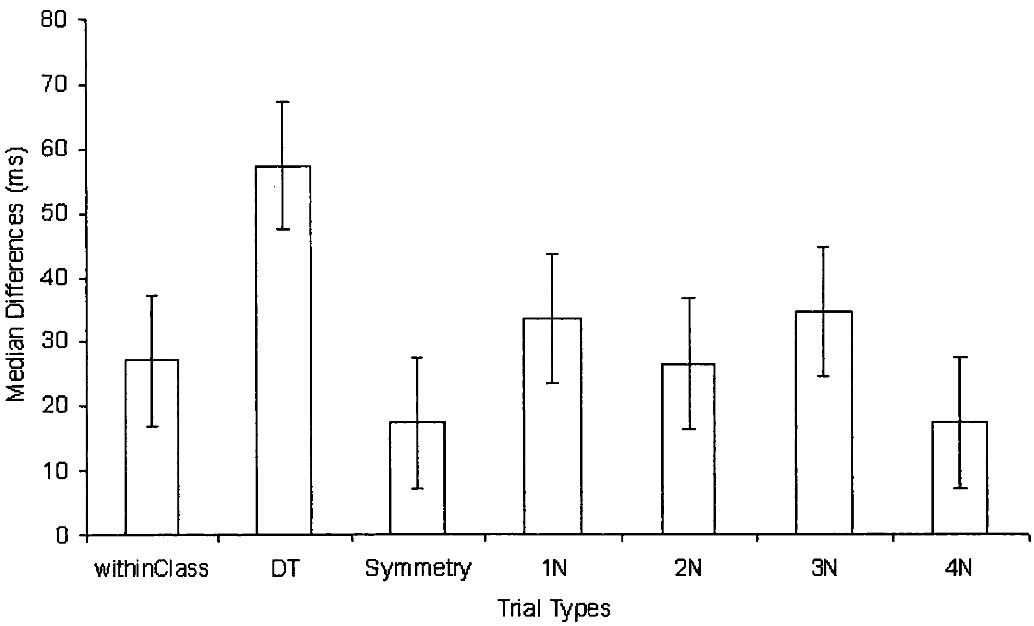


Figure 24: Median of mean differences between NoClass and WithinClass/DT/Symmetry/1N/2N/3N/4N, and 95% confidence intervals across all passed participants

The averaged waveforms generated by NoClass trials (e.g. A1A3) are almost identical to the waveforms generated by class inconsistent trials (i.e.. CrossClass, A1A2), therefore, only CrossClass trials served as a comparison to the directly trained and equivalence related trials in the statistical analysis. It was also found a large positivity around 300 ms after target onsets in frontal area, therefore, statistical analyses were performed in the two timeframes: 250 – 350 ms, and 350 – 550 ms.

5.3.2.1.1 250 – 350 milliseconds

In 250 – 350 ms timeframe, three repeated measures of ANOVAs with trial type (directly trained, equivalence, and CrossClass, 1-node, 4-node), electrode sites (F3-4-Z, FC1-2, C3-4-Z) as repeated factors, were calculated in 3 frequency range: Theta (4-8 Hz), Alpha (8-13 Hz), and Beta (13-22 Hz), respectively. Significant main effects were found from 8 – 13 Hz (Alpha frequency) in trial type, $F(4, 120) = 3.105$, $p=.018$; sites, $F(7, 210) = 21.654$, $p<.0001$. Further analysis showed no significant differences between trial types. A significant interaction was also found between trial type and sites, $F(28, 840) = 1.581$, $p=.029$. Bonferroni pairwise tests were performed subsequently, see Table 31 for a summary of the significant results obtained. No significance was found in Theta and Beta frequencies in this timeframe.

Table 31: Results of Bonferroni pairwise tests comparing direct trained (DT), equivalent (EQ), and CrossClass (CC), 1-node (1N), 4-node (4N), trials from 250 – 350 ms in 8 – 13 Hz.

	Significant results				
	Left sites	Middle sites		Right sites	
Pairwise	C3	FZ	CZ	C4	FC2
DT vs. EQ	a				
EQ vs. CC	a				
1N vs. 4N		a	a	a	a

[a] p<.05.

5.3.2.1.2. 350 – 550 milliseconds

For the 350 – 550 ms timeframe, three repeated measures of ANOVAs with trial type (directly trained, equivalent, and CrossClass, 1-node, 4-node), electrode sites (F3-4, FC1-2, C3-4, CP1-2, P3-4-Z, O1-2-Z) as repeated factors, were calculated in 3 frequency range: Theta (4-8 Hz), Alpha (8-13 Hz), and Beta (13-22 Hz), respectively. Significant main effect were found from 8 – 13 Hz (Alpha frequency) in sites, $F(15, 450) = 25.791$, $p<.0001$. Significant interaction was also found between trial type and sites, $F(60, 1800) = 1.489$, $p=.01$. Subsequent pairwise tests demonstrated significant differences between 1-node and 4-node trial type, between direct trained and equivalent trials; were approaching significance between equivalent and class inconsistent trials. Results of these tests were shown in Table 32. Approaching significance was found amongst trial types, $F(4, 120) = 2.131$, $p=.081$; But no significant difference between trial types. Significant main effect were also found from 13 – 22 Hz (Beta frequency) in trial type, $F(4, 120) = 4.852$, $p=.001$; sites, $F(15, 450) = 13.641$, $p<.0001$. However, no significant results emerged from the bonferroni pairwise comparisons between trial types. There was no significance for the Theta frequency in this timeframe.

Table 32: Results of Bonferroni pairwise tests comparing direct trained (DT), equivalent (EQ), and CrossClass (CC), 1-node (1N), 4-node (4N), trials from 350 – 550 ms in 8 – 13 Hz.

	Significant results					
	Left sites	Middle sites			Right sites	
Pairwise	CP1	CZ	PZ	OZ	CP2	FC2
DT vs. EQ			[a]			
EQ vs. CC	.071	.065	.074		.085	
1N vs. 4N		[a]		[a]	[a]	.054

[a] p<.05.

In summary, priming effects were significant across all trial types, except 4-node trials. The largest priming effect was found in directly trained trials, the differences between certain trial types were approaching significance (i.e., DT vs. Symmetry; 1N vs. 2N). Analysis of ERPs indicated that significantly greater positivity in 4-node trials compare to 1-node trials around 250 to 350 ms after target onset on middle and right hemisphere sites (See Figure 25). Significantly greater positivity in equivalence trials than class inconsistent trials and direct trained trials was found around 250 to 350 ms after target onset on site C3 (See Figure 26). Moreover, significantly greater negativity in 4-node trials than 1-node trials was found around 350 to 550 ms after target onset on middle and right hemisphere sites (See Figure 27). Significantly greater negativity in equivalent trials than directly trained trials was found around 350 to 550 ms after target onset on site PZ. Differences between equivalent and class inconsistent trials were found approaching significance on left, middle and right hemisphere sites.

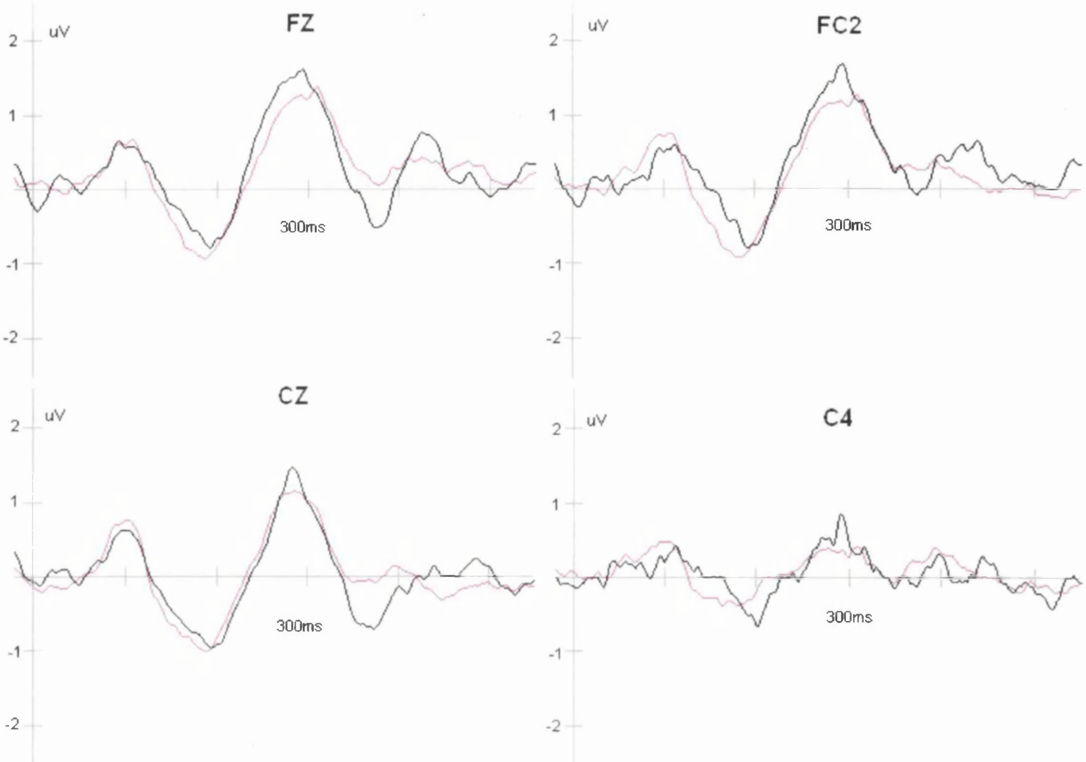


Figure 25: Grand average waveforms across all passed participants in sites FZ, CZ, FC2, and C4. The black line indicates 4-node trials, the red line indicates 1-node trials.

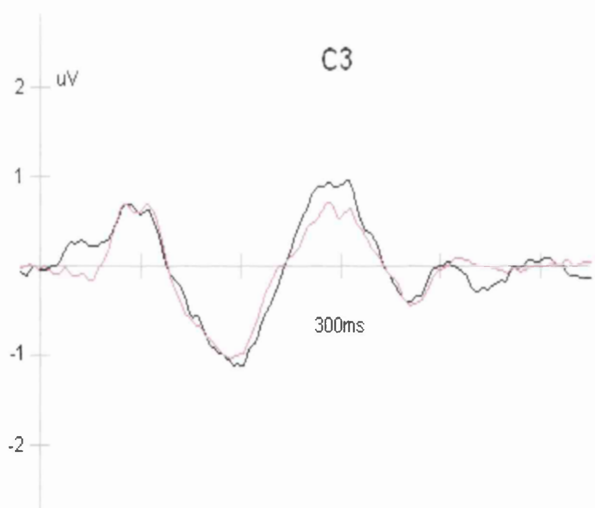


Figure 26: Grand average waveforms across all passed participants in site C3. The black line indicates equivalence trials, the red line indicates class inconsistent trials.

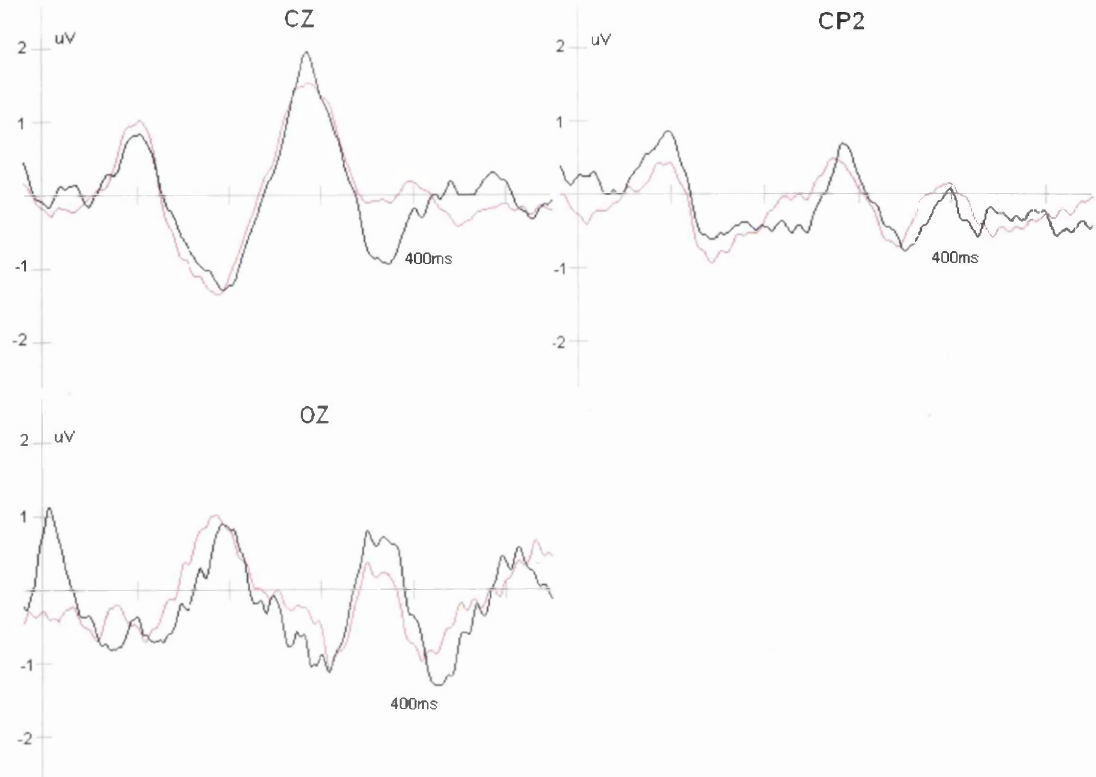


Figure 27: Grand average waveforms across all passed participants in sites CZ, OZ, and CP2. The black line indicates 4-node trials, the red line indicates 1-node trials.

5.3.2.2. Failed Participants

10 participants failed to reach the mastery criterion of 80% above accuracy in the equivalence test. There was difference in RTs between NoClass trials and WithinClass trials, but this difference was not statistically significant.

5.3.3 Discussion

The analysis of RTs for the passed participants suggested that the strength of priming was highly correlated with experimental training as the RTs in the directly trained trials demonstrated the largest priming across trial types. This finding is in line with the emergent ERPs in Experiment 7. Although priming effects were found across all node relations, they did not follow a liner pattern in accordance with nodal number. This confirmed the speculation that nodal number effects are not necessarily linear relations (Sidman, 1994). However, the priming effect demonstrated for the 4-node relations was the only one that did not reach statistical significance when compared to the other node relations (i.e.. 1, 2, and 3-node), this lack of significance seems partially consistent with the assumption of nodal numbers, that is, the strength of priming is greater in stimulus pairs that have less node intercepts between them, whereas it is smaller in stimulus pairs that have more node intercepts in between them (Fields & Moss, 2007). Again, the lack of priming in failed participants replicated the findings of Experiment 7 and previous studies (Barnes-Holmes, 2005 Exp2).

Analysis of mean amplitudes for the passed participants demonstrated a greater positivity 300 ms after the target stimulus was presented in 4-node than 1-node trials, in equivalence trials compared to directly trained trials and class inconsistent trials. These findings clearly resembled the P300 commonly associated with categorization. Another characteristic of the P300 is to detect the level of effort participants had put into the task (Johnson, 1984, 1986, 1988). Specifically, the P300 is greater when participants devote more effort to a task. This has led researchers (Isreal, et al., 1980) to argue that P300 can be used to detect source allocation. Logically, as nodal number increases, the more effort is expected to be devoted to the task. Therefore, P300 positivity should be larger in 4-node than 1-node trials, the current data sets confirms this speculation in frontal, central lobes and also confirmed right hemisphere advantages in reasoning and decision making. Unfortunately, the current study failed to replicate the greater negativity in N400 in class inconsistent trials when compared to equivalence trials, even though there were few sites approaching significance in the central and parietal lobes. Interestingly, a greater N400 was found in 4-node rather than 1-node trials in central, parietal and occipital lobes. This can not be due to some uncontrolled event, since it was consistent with the lack of priming in 4-node trials only when measuring RTs across trial types. This consistency between the results from participants RTs and ERPs is in line with the nodal number hypothesis, as nodal number increases, stimulus control decreases, therefore, resulting in the weakest priming and largest emergence of mismatching in 4-node relations than 1-node relations.

5.4 General Discussion

The aim of the current chapter was to pinpoint the neurological processes underlying the formation of equivalence classes and nodal distance between class members. Experiment 7 employed a 2x2 design that systematically manipulated reinforcement and number of trial presentations. 39 participants established two four-member equivalence classes using a standard MTS paradigm. Experiment 8 employed a simple within subject design in which 31 participants established two six-member equivalence classes using a standard MTS with a serial training structure. Both experiments employed a lexical decision task between baseline discrimination training and tests for equivalence. The different nature of the experimental designs employed renders it difficult to compare the results from the two experiments. However, several broad findings were consistent across both.

The major consistent finding across experiments was the lack of robust N400 in class inconsistent trials in both experiments. This finding is consistent with Haimson, et al. (2009)'s Experiment 2, in which, the N400 effect was only observed in participants who had received the electrophysiological testing after equivalence testing, but not before. As Experiment 7 and 8 both employed electrophysiological testing before equivalence testing, the lack of robust N400 effect might be due to the lack of equivalence testing which provided a critical context control for the emergence of equivalence relations (Sidman, 1985). The two studies reported herein not only replicated the findings of Haimson, et al.'s (2009) Experiment 2, but also extended the finding from 8 participants to 70 participants. However, this finding is contradictory to Barnes-Holmes, et al.'s (2005) Experiment 3. These authors argued that mediated priming effects can only be assessed when an equivalence test is not presented before the electrophysiological testing. Due to the fact that class consistent stimulus pairs were presented repeatedly during the equivalence testing. It is difficult to define the priming effect observed as "direct priming" or "mediated priming". This is fatal when assuming equivalence (Sidman, 1994, 2000) or derived stimulus relations (Hayes, et al., 2001) as a behavioural account of semantic network formation. Of course, there might be other procedural differences that could account for the lack of N400 effects in Experiment 7 and 8. For example, Barnes-Holmes, et al.'s (2005) Experiment 3 used two four-member equivalence classes whereas Yorio, et al., (2008) employed two three-member equivalence classes, both with a within subjects design. However, Experiment 7 used a 2x2 between subject design that manipulated reinforcement and trial presentation these variables may have affected the formation of equivalence classes and masked the N400 effect with direct learning being imposed by reinforcement and trial presentation. Hence, only directly trained trials produced the N400 (Experiment 7) rather than equivalence trials and class inconsistent trials. Experiment 8 employed a within subject design but with extended equivalence classes, identical to that used in Haimson, et al. (2009). Specifically, it involved two six-member equivalence classes. The analysis of mean amplitude suggested that the N400 effect was masked by the control of nodal number in extended equivalence classes.

Another consistent finding between the two experiments was that mean amplitudes demonstrated inflated positivity 300 ms after the target onsets, which is in line with the findings in Yoiro, et al., (2008). This finding reflected the P300 that was commonly associated with source allocation, categorization, and decision making (Azizian, Freitas, Watson, & Squires, 2006). However, there is a difference in the P300 reported between Experiment 7 and 8. In Experiment 7, the P300 seems associated with the level of effort participants put into the task, the difficulty of the task was influenced by reinforcement and trial presentation. Whereas in Experiment 8, the P300 seems to be associated with categorization and the level of effort participants put into the task, which is in line with the assumptions of nodal number effect.

Both experiments confirmed that directly trained trials were the most sensitive towards the effects of experimental manipulations on measures of RTs and ERPs (Experiment 7), when compared to equivalence and class inconsistent trials, providing evidence for the importance of empirical training on the formation of semantic networks. Only participants who formed equivalence classes demonstrated priming effects in both experiments, the lack of priming effects in participants who did not form equivalence is in line with previous studies (Barnes-Holmes, et al., 2005 Exp2), and supported the assumption that equivalence (Sidman 1994, 2000) or derived stimulus relations (Hayes, et al., 2001) provide a behavioural account for semantic network formation. Both experiments reported major ERP changes based on experimental manipulations during 8 – 13 Hz frequency range, however, there were no or little changes from 4 – 8 Hz, 13 – 22 Hz, this finding also suggested that the experimental manipulation seems more sensitive in the alpha frequency than in the beta or theta frequency.

Additionally, a greater N400 was found for 4-node rather than 1-node trials in the central, parietal and occipital lobes in Experiment 8. This cannot be attributed to some uncontrolled event, since it was consistent with the lack of priming on 4-node trials only when measuring RTs across trial types in the same experiment. This consistency seems in line with the nodal number hypothesis, as nodal number increases, the stimulus control decreases, therefore, results in the weakest priming and greater sense of mismatching in 4-node relations than 1-node relations.

In conclusion, the lack of N400 effects from the 8 to 13 Hz frequency range in Experiment 7 and 8 is in line with Haimson, et al., (2009) and Sidman's (1985) account of equivalence, however, contradicts Barnes-Holmes, et al., (2005)'s Experiment 3. These contrasting findings might be due to the different experimental procedures employed. Findings from Experiment 7 and 8 also suggested that the P300 and N400 are neural markers for the nodal number effect, reinforcement and trial presentation.

Chapter 6: General Discussion

The current chapter aims to provide the reader with a brief summary of the empirical work presented in Chapters 2–5. Having summarised the work from Chapter 2 and 3, some of the relevant theoretical issues arising from these two chapters will be considered. This will be followed by a summary and discussion of the relevant theoretical issues in Chapter 4. Finally, a summary of Chapter 5 will be presented and the current chapter will close with a concluding commentary.

6.1 Chapter 2 and 3: Summary

The first two chapters of the current work aimed to investigate whether nodal distance in the equivalence class literature is a genuine effect or an experimental artefact (Imam, 2001, 2006). To that end, Chapter 2 comprised of four experiments that focused on systematically manipulating training structure and reinforcement history during conditional discrimination training. Experiment 1 employed a 2x2 between-subject design, with training structure (serial versus simultaneous) and time of trial presentation (limited hold versus no limited hold) as independent variables. The combination of these two variables resulted in four experimental conditions: Equal Reinforcement LH (limited hold), Equal Reinforcement no LH, Unequal Reinforcement LH, and Unequal Reinforcement no LH. In the Equal Reinforcement conditions, the number of delivered reinforcers was approximately equal across all baseline trial types, and all trial types were introduced simultaneously. In contrast, in the Unequal Reinforcement group, the number of delivered reinforcers differed across trial types, and trial types were introduced serially. If nodality is a result of unequal reinforcement history, it would only emerge under unequal reinforcement conditions, whereas if time contingency is a crucial in facilitating this emergence, nodality should only emerge under the Unequal Reinforcement LH condition. Two five-member equivalence classes were trained and tested using a standard MTS paradigm. 15/21 university students formed equivalence at 90% accuracy after repeated training and testing. This resulted in 4 participants in each condition, with the exception of the Equal Reinforcement LH condition in which only 3 participants formed equivalence. The results indicated that mean RSs for correct responses from participants who formed equivalence was an inverted function of nodal number only for the Equal Reinforcement no LH condition. Correct responses from participants who formed equivalence successfully were an inverted function of nodal number in certain test blocks across conditions that had no time contingency in place (no LH conditions). No nodality effect was found for participants who did not form equivalence. The overall findings of Experiment 1 supported the assumption that nodality is not a result of unequal reinforcement during conditional discrimination training. Additionally, the emergence of nodality in the no LH conditions might also explain Imam's failure to demonstrate nodality when a time-accuracy trade off occurs (Imam 2001, 2006).

No nodal effect was only found under equal reinforcement without limited hold, which suggested that nodal effects are not a result of unequal reinforcement during conditional discrimination training. Experiment 2 aimed to replicate the Equal Reinforcement no LH condition in order to investigate whether the positive results obtained in Experiment 1 could generalize to a larger sample. 12 out of 23 participants formed equivalence after repeated training and testing cycles. Unfortunately, no sound nodal number effect was observed in response time and response accuracy across either the participants who passed or failed equivalence testing, although there was a decrease in mean RSs as nodal number increased across all passed participants.

The contradictory results obtained in the first two experiments raised a number of questions. First, the lack of nodal number effects in the larger sample of the Equal Reinforcement no LH condition may have been a result of large inter-participant variability that may have reduced the power of the analysis of averaged responses. Second, impaired baseline relations in the test for equivalence was greater in Experiment 2 than in Experiment 1, the impaired baseline relations and the low equivalence pass rate might have jeopardised the emergence of nodal number effects. According to Fields et al., (1993, 1995, 2007, 2008) the standard MTS test paradigm maximizes class based rather than nodal number based discriminations, thus nodal effects often disappear after repeated training and testing cycles. Whereas transfer of function tests are reported to maximize nodal based discriminations, and maintain nodal number effects even after repeated training and testing cycles. Therefore, Experiment 3 and 4 sought to investigate the maintenance of nodal number effects under equal reinforcement. Experiment 3 replicated the Equal and Unequal Reinforcement no LH conditions in Experiment 1, with an additional transfer of function (Tof) paradigm adapted from Fields et al. (1993, 1995, 2008). The Tof paradigm was included at the end of the MTS test to examine whether the amount of responses transferred amongst equivalent stimuli was a function of nodal numbers. In the function-training phase, differential responses were trained to the C and D stimuli in each class using corrective feedback. Next, the A, B, E, and F stimuli were presented in the absence of corrective feedback and the number of responses to each stimulus was observed. If the prediction that equal reinforcement eliminates a nodal effects is correct, then responses to A, B, E and F should be distributed equally following Equal Reinforcement training. In contrast, if the nodal account is correct, then the A and B stimuli should evoke the response trained to the C stimulus, and the E and F stimuli should evoke the response trained to the D stimulus, despite the differential reinforcement. In addition, if unequal reinforcement is indeed a confounding variable then, following unequal reinforcement training, responses trained to the C stimulus should transfer to the A and B stimuli more readily than responses to the D stimulus transfer to the E and F stimuli. That is, B-C and A-B relations should be more strongly established, whereas D-E and E-F relations should be weak. Eight participants were randomly assigned to either the Equal or Unequal Reinforcement conditions. Two six-member equivalence classes were established after repeated MTS training and testing. Participants were required to complete 10 consecutively correct trials before proceeding to the test phase in order to stabilize baseline performance. Typical measures of RS and correct responding suggested that mean RSs were not an inverted function of nodal number under either Equal or Unequal Reinforcement, correct responses were varied unsystematically across node-related trial types, with one exception (i.e., participant) in each condition who demonstrated nodality on certain test blocks. This might suggest correct responses were disrupted as the equivalence class was extended from five (Experiment 1 and 2) to six (Experiment 3) members (Sidman, 1994; Fields et al., 1997, 2007).

Response transfer was assessed by measuring the relative frequency with which the responses trained to the C and D stimuli were evoked by the other experimental stimuli. For example, the C1 stimulus was an S^D for 3 spacebar presses; if 3 spacebar presses were evoked during all six presentations of the A1 stimulus then the relative frequency was 100%. In some cases, the sum of responses trained to the A, B, E or F stimuli in a class did not equal 100%. This occurred because some responses other than those trained to the C and D stimuli were evoked (e.g. across-class errors). If the response trained to a C or a D stimulus was evoked by other members of the same class, then the response function was deemed to have transferred in accordance with an equivalence relation. More importantly, if the response trained to C was evoked by the A and B stimuli, and not by the E and F stimuli, within the same equivalence class, then those responses were deemed to be also under the control of nodal number (similarly, if the response trained to D was evoked by the E and F stimuli and not by the A and B stimuli). Fisher's exact tests suggested that response transfer was a function of nodal number for the members of one of the equivalence classes (e.g. class A1, B1, C1, D1, E1, and F1) in both the Equal and Unequal Reinforcement conditions. Overall the findings of Experiment 3 suggested that differential reinforcement and a serial training protocol resulted in differential response transfer. Transfer from the D1 stimulus to the E1 and F1 stimuli was as robust as the transfer from the C1 stimulus to the A1 and B1 stimuli. In the Equal Reinforcement condition, the untrained responses that transferred from the C2 and D2 stimuli were under the control of nodal number; hence unequal reinforcement was not a prerequisite for nodal number effects to emerge.

Experiment 4 replicated the Equal Reinforcement condition in Experiment 3, with an extended sample size, Extinction of baseline relations in testing for equivalence in Experiment 3 suggested an increase in the number training trials was necessary. Therefore, 20 consecutively correct trials were required before proceeding to the test phase. Results obtained from mean RSs replicated the nodal number effects seen in Experiment 1. Moreover, the proportion of response transfers was a function of nodal number under equal reinforcement and simultaneous training, which is consistent with the findings from Experiment 3. The strength of node effects obtained in Experiment 4 increased in both measures after stabilized baseline simple discriminations were demonstrated during testing.

Results from the four empirical studies in Chapter 2 suggested that nodal number was a predictor of RS even when reinforcement for responding to baseline trial types during conditional discrimination training was equalized. In Experiments 3 and 4, a transfer of function test was employed as an additional measure to examine nodal effects. Again, nodal effects were observed in the Equal Reinforcement condition. Furthermore, a serial training protocol that resulted in unequal reinforcement did not appear to influence nodal number. That is, less reinforcement to particular trial types did not result in poorer transfer.

Experiment 5 in Chapter 3 aimed to investigate the delayed emergence of equivalence from unequal frequency of stimulus presentations during MTS training. This discrimination account proposed by Saunders and Green (1999) argued that the delayed emergence of equivalence classes was a result of differential reinforcement from unequal presentation of the stimuli during MTS training, rather than node number based differential control over stimulus pairs (Fields, et al., 1993, 1995, 2008; Kennedy, 1991; Kennedy, et al., 1994; Bentall, et. al., 1993). Experiment 5 aimed to compare and contrast the predictions of both the discrimination account suggested by Saunders and Green (1999), and the nodal account, in order to determine which set of predictions could better account for delayed emergence of relation types in equivalence class formation. In this study, participants were trained and tested across two classes of five stimulus members on an MTS procedure with a simultaneous training protocol resulting in equal trial presentations on each trial type. A gradual feedback fading procedure was employed with feedback fading from 100%, followed by 50%, to 0% across training trials, as participants' performance can be impaired by the sudden elimination of feedback in the test phase. In order to stabilize performance, 16 consecutively correct responses across training trials with 0% feedback was required before proceeding to the equivalence test phase.

In the case of a five member equivalence class, according to the nodal account equivalence class formation emerges as a function of nodal number, that is, 1-node relations (AC, CA, BD, DB, CE, EC) emerge first, followed with 2-node relations (AD, DA, BE, EB), with 3-node relations (AE, and EA) emerging last. According to the discrimination account the formation of equivalence is a function of discriminations acquired during baseline training; thus, stage 1 relations (BD, and DB) should emerge first, followed by stage 2 relations (AC, CA, CE, EC, AD, DA, BE, and EB), while stage 3 relations (AE and EA) emerge last. Because of the overlap in assumptions of the two accounts, it is necessary to separate the effects by comparing performance between node relations and stage relations, while keeping other variables constant (i.e., training structure, reinforcement, number of trial presentation). Therefore, if the nodal account is correct, responses to stage 2 and also 1-node relation types should be faster and more accurate than responses to stage 2 and also 2-node relation types. If the discrimination account is correct, responses to stage 1 and also 1-node relation types should be faster and more accurate than responses to the stage 2 and also 1-node relation types.

21/47 participants formed equivalence, in which, 15 of them reached the mastery criterion of 85% correct immediately after conditional discrimination training. Wilcoxon signed-rank tests indicated that response speed from participants who formed equivalence was significantly faster on the stage 2 and also 1-node relation type than on the SG2/2N relation type. Correct responses for participants who formed equivalence demonstrated 30% of cycles were consistent with the nodal account; 20 % of cycles were consistent with the discrimination account and 5% of cycles were consistent with both accounts. Participants reached ceiling performances on all three relation types on 11% of cycles. The remainder of cycles were inconsistent with either account. Moreover, 29% of passed participants showed sole consistency with the nodal account, 24% were solely consistent with the discrimination account, and 33% of passed participants demonstrated features of both accounts in different cycles, 14 % of passed participants showed features of neither account. Correct responses in participants who did not form equivalence demonstrated 33% of cycles were consistent with the discrimination account; 26 % of cycles were consistent with the nodal account; only 1% of cycles were consistent with both accounts. The remainder of cycles were inconsistent with either account. Moreover, 40% of the failed participants showed sole consistency with the discrimination account, 32% were solely consistent with the nodal account, and 16% of participants demonstrated features of both accounts in different cycles, 12% of participants showed features of neither account. Overall the findings of Experiment 5 in Chapter 3 was that averaged RSs were significantly faster in 1-node than 2-node relation types, when stimulus frequency was kept constant (stage 2 relation type). Correct responses across participants who formed equivalence after repeated exposure to training and testing cycles again favoured the nodal account, as there were more nodal consistent cycles than discrimination consistent cycles (22 vs. 15). There is little to suggest that the discrimination account is an accurate predictor of delayed emergence, even though there was higher yield for the discrimination account in correct responses from people who did not form equivalence. This might suggest that the emergence of nodal relations has a facilitative role in overall equivalence class formation.

6.2 Theoretical Issues

Chapter 2 and 3 contained a series of studies investigating the delayed emergence of equivalence during a standard MTS paradigm. The findings of these studies are in accordance with previous research by Fields and his colleagues (Fields & Watanabe-Rose, 2008; Fields & Moss, 2007), whereas conflict with research suggesting differential reinforcement during baseline training was responsible for delayed emergence of equivalence performance (Imam, 2001, 2006; Sidman 1994; Saunders and Green, 1999). The four studies in Chapter 2 highlighted the importance of the test format when examining the relatedness of stimuli in equivalence classes. Three measures were employed in the current study: response accuracy, RS, and transfer of function. Response accuracy was not significantly different in any of the conditions. Response speed was significantly different in Experiments 1 (in the Equal Reinforcement condition) and in Experiment 4. It was the transfer of function test, however, that demonstrated the clearest evidence of nodal effects for the Equal Reinforcement group. Fields and Watanabe-Rose (2008) have speculated that the format of the MTS paradigm itself occasions responding in accordance with class membership and discrimination *between* classes. In contrast, the format of the transfer of function test is such that responses occur in the presence of members of the same class and therefore occasions responding according to *within*-class differences, such as nodal number. Interestingly, RS, although measured during MTS trials, appears to vary as a function of nodal number, although the sensitivity of RS is less than that of the transfer of function test. The importance of these studies not only lends support to the nodality literature, but also generalized the results to a bigger sample group. The number of participants who demonstrated equivalence under equal reinforcement (n=19) is substantially greater than in previous studies (two participants in Fields et al., 1995 and four participants in Fields and Watanabe-Rose, 2008). As stated by Fields and Watanabe-Rose's (2008) "additional research will be needed to determine whether a larger segment of the population would also bifurcate class membership based on nodal structure" (p. 378)

According to Fields and Moss (2007), the test format in Chapter 3 emphasised the between-class discrimination in different class sets, which theoretically maximized the between-class discrimination control, this might account for the absence of nodality in some test cycles for some participants. However, even in light of this difficulty, more than half (13/21) of the participants who were successfully trained in equivalence (after repeated exposure to the training and testing cycles) demonstrated nodal patterns in their correct responses on certain cycles. This again, supported the reinforcement contingency theory of equivalence formation (Sidman, 1994) extended by Fields and Moss (2007), that is, test performance is contingent upon reinforcement history, and that the test format emphasis on either between-class discrimination in different class sets or class-based relations (i.e.. nodality) in the same class was responsible for the expression of nodal effects in some studies, but not others. They further argued that the contingency upon test format employed specifies contextual cue that controls subsequent stimulus relations, therefore, in line with the contextual control hypothesis proposed by Sidman's equivalence and Relational Frame Theory (Hayes, et al., 2001).

There were some differences between those studies presented in the first two chapters and the previous studies that varied the training protocol and level of reinforcement (Imam, 2001, 2003, 2006). Similar to many other equivalence studies (Fields et al., 1993; Barnes-Holmes et al., 2005) the current studies employed two, rather than three, comparison stimuli. The effects of nodal number have been shown to diminish with the addition of a third stimulus group (Kennedy, 1991 Exp. 2), and thus the inclusion of a third stimulus may have made interpretation difficult if nodal effects were not observed in Experiment 2 . Although studies have demonstrated nodal effects with various types of stimuli (e.g., pictures, letters, symbols), the present studies employed pronounceable letter strings, whereas Imam (2001, 2003, 2006) used graphical stimuli. Is it possible, indeed likely, that relations among stimuli were rapidly learned because pronounceable, rather than graphical, stimuli were employed (cf. Imam, 2001, 2003, 2006). However, the use of these stimuli did not seem to facilitate learning to such a degree that the differential reinforcement delivered between Unequal and Equal groups was attenuated, rather the training protocol *per se* and numbers of baseline training counts (Chapter 2).

Imam (2001, 2003, 2006) and Spencer and Chase (1996), included only those trial types involving either the most trained conditional relations or the least trained conditional relations in their analyses. According to Spencer and Chase (1996), this exclusion procedure was designed partly to account for imbalances in "the order and differential amount of training on each baseline conditional discrimination" (p. 649). However, the aim of the studies presented herein was to measure the influence of these factors, and therefore it seemed counter-intuitive to exclude these trial types. Saunders and Green (1999, p. 132) have also noted that only including the most and least trained stimuli in the analyses introduces some additional problems in interpretation.

As the first empirical attempt to compare the nodal account with discrimination account in explaining the delayed emergence of relation types in equivalence class formation, the analysis of correct responses outlined in the methods section in Chapter 3 provided a useful example of how to approach the two theories (nodal and discrimination accounts) within the same data set. One might argue that the nature of comparisons is not mutually exclusive, responses likely fall into a linear pattern in which participants demonstrate better responding on SG1/1N compared to SG2/1N, followed by SG2/2N relations. However, this was not the case in the Experiment 5, as participants demonstrated such a linear pattern on only 4/74 of the cycles, which is way beyond chance level.

6.3 Chapter 4: Summary

Chapter 4 aimed to investigate the relative contribution and interaction between reinforcement, time of acquisition and number of stimulus presentations in forming and maintaining connections among nodes in a semantic network. 249 university students were trained and tested on two four-member equivalence classes using a standard MTS paradigm. In the training phase, participants were randomly assigned to eight experimental conditions that were systematically manipulated in terms of reinforcement (100% versus 50%), time of acquisition (Phased versus Simultaneous), and no. of stimulus presentations (Unequal versus Equal). After a fixed number of training trials (180), each participant was presented with a lexical decision task, similar to the paired-association paradigm discussed in Chapter 1 (Section 1.2.2.1). In which, baseline relations and all derived relations were presented successively with a mixture of novel nonsense word pairs in a single block without feedback. Participants were required to press the "Yes" key if the second word in a given pair had been presented in the baseline training phase, and to press the "No" key if the second word in a given pair was a novel nonsense word. After a fixed number of trials (158), a MTS test for baseline relations and derived relations was presented in a single randomized block.

102 participants formed the two equivalence classes at 70% above accuracy immediately after conditional discrimination training. The success of equivalence performance was lower than 50% of participants across conditions, with the exception of the unequal no. of stimulus presentations condition when reinforcement and time of acquisition was held constant. RTs obtained from the lexical decision task indicated that participants responded faster to equivalence related stimulus pairs than non equivalence related stimulus pairs in all 100% reinforcement conditions. Significant main effects were found between RTs from related pairs and pairs from different equivalence classes, amongst conditions, and a significant three way interaction was also found in a mixed analysis of variance. However, subsequent tests showed RTs were only significantly faster in equivalence related stimulus pairs compared to non equivalence stimulus related pairs in the 100% reinforcement condition when stimulus pairs were learned early and no. of stimulus presentations was held equal. Participants who did not form equivalence showed some signs of priming effects, but this was not statistically significance. Overall the findings of Experiment 6 presented in Chapter 4 suggested that level of reinforcement is the single most salient contributor in determining the strength of connection of new words joining an existing semantic network when compared to time of acquisition and no. of stimulus presentations. Interestingly, the latter are the most commonly reported contributors in the cognitive literature.

6.4 Theoretical Issues

The study reported in Chapter 4 supported the previous findings of Barnes-Holmes, et al 2005's Exp2, in that participants RTs were faster for related stimulus pairs than pairs that were from different equivalence classes, only in cases where participants formed equivalence at the end of the experiment. The differences found between passed and failed participants again supported the argument that derived relations, rather than directly reinforced stimulus relations alone, provide a behavioural model of semantic networks (Barnes-Holmes, et al 2005; Barsalou, 1999; Deacon, 1997; Hayes & Bisset, 1998).

All priming effects that were observed emerged during conditions where reinforcement levels were kept relatively high (100%), regardless of the other two factors. These findings suggest that reinforcement alone as studied in the behavioural literature is sufficient to produce what cognitive researchers refer to as semantic network growth (Imam 2001, 2003, and 2006). Additional support for this postulate emerged from the analysis of cross condition interactions, that is, the priming effect was significantly larger when reinforcement was kept at 100%, even when the other two factors were held constant. Thus indicating that reinforcement is the most important single factor in predicting the connective strength between a newly introduced word and the existing network, which might mask any effects of when and how frequently the word was encountered. These findings coincide with Sidman's reinforcement contingency theory, that is, reinforcement contingency is the only major contributing factor on a simple discrimination preparation that results in between-class equivalence performance. Thus, the results further confirmed the postulate that equivalence (Sidman, 1994) or derived relations (Hayes, et al, 2001) may usefully account for the primary component unit in a semantic network.

The observed priming effect in Condition 7 (50%ReinfEquPresentPhasedA), but not in Condition 8 (50%ReinfEquPresentSimultA), also suggested words learned early had stronger connections than late when reinforcement and no. of stimulus presentations were under control. Moreover, across 8 conditions, Unequal no. of stimulus presentations condition yielded the highest pass rate (71%); while stimulus pairs that were 50% in reinforcement, equal in no. of stimulus presentations and simultaneously in acquisition produced the lowest yield (31%), despite the largest sample size (65) amongst conditions. This might suggest that the effect of no. of stimulus presentations is not strong enough to express itself in lexical decision task, even when reinforcement and time of acquisition were controlled for, but can be expressed subtly in a test for derived stimulus relations. These findings suggested that participant number has little to contribute to the ratio of success of equivalence performance in 50%ReinfEquPresentSimultA condition.

Although statistically significant main effects were observed between equivalence pairs compared to non equivalence pairs, across the 8 conditions, and there were interactions between trial type and condition, follow-up tests demonstrated far less significances between specific variables than predicted. This lack of significance in follow-up tests, particularly, in the 100% ReinfUneqPresentPhasedA condition, might be due to reduced power in Experiment 6. First, fewer participants formed equivalence in every condition and more than half of them failed to form equivalence after a lexical decision task. Second, the low (70%+) accuracy criterion on the equivalence test restricted stable response to emerge.

Despite the manipulation of reinforcement, no. of stimulus presentations and time of acquisition, the low pass rate in the Experiment 6 may also be a result of a different procedure than that employed by Barnes-Holmes's Exp2. This experiment also applied a fixed number of training (180) trials in terms of those three factors during the equivalence training phase, unlike in Barnes-Holmes's Exp2, in which 24 consecutively correct trials were required before completing one training exposure out of 10. As a result of this intense training, their participants were exposed to more baseline relations (413 – 424 trials) than participants in the current study, thus, their baseline relations were more firmly established prior to the derived relations test, this might explain the high failure rate in the Experiment 6.

There were other differences between the two studies. There was a longer SOA (600ms) in the lexical decision task in the current study, with the same 200ms latency of the prime stimulus as in Barnes-Holmes's Exp2, the blank screen between the prime and the target was increased from 50ms to 400ms in the current study. Despite the different SOAs used, these would all be considered as short SOAs in the cognitive literature (Neely, 1991). Moreover, eight instead of six novel nonsense words were used in the current lexical decision task, with balanced trial types in the presentation of stimulus pairs from each different equivalence class, which resulted in an increased overall number of trials being presented. In addition, each trial type was presented only once in the current lexical decision task whereas each was presented twice in Barnes-Holmes's Exp2.

6.5 Chapter 5: Summary

The two experiments reported in Chapter 5 aimed to pinpoint the neurological processes underlying levels of reinforcement, number of trial presentation, and nodal numbers in equivalence class formation. The aim of the first experiment was to investigate the neurological processes underlying reinforcement and trial presentation in equivalence formation. In which, levels of reinforcement (high versus low) and numbers of trial presentation (high versus low) were manipulated resulting in four experimental conditions (HRLowTrial, HRHighTrial, ERLowTrial, ERHighTrial). The procedure is similar to that employed in Chapter 4, with EEG as additional measure recorded during the lexical decision task. 39/58 participants formed two four-member equivalence classes at 80% or above accuracy across four conditions. RTs were faster for related stimulus pairs than pairs from different equivalence classes across all conditions, except the ERLowTrial condition. Statistical analyses showed that none of these differences reached significance.

Mean amplitude was analyzed across two timeframes: from 250 to 350 milliseconds, and from 350 – 550 milliseconds. In the 250 – 350 milliseconds analysis the mean amplitudes from 8 to 13 Hz frequency range demonstrated significant differences between the HRHighTrial and HRLowTrial conditions, between the HRLowTrial and ERLowTrial conditions on directly trained trials. Significant differences also emerged between the HRHighTrial and HRLowTrail conditions in class inconsistent trials. In the 350 – 550 milliseconds analysis the mean amplitudes from 4 – 8 Hz demonstrated significant differences between class inconsistent trials and stimulus equivalence (i.e., derived) trials regardless of experimental manipulation. The mean amplitudes from the 8 to 13 Hz frequency range demonstrated significant differences between the HRLowTrial and ERLowTrial conditions on the sites O1, and OZ on directly trained trials and was approaching significance at sites C4 and FC2. Right hemisphere sensitivity to the effects of the experimental manipulations on directly trained relations was also observed. No differences were found on RTs of the participants who did not form equivalence.

Experiment 8 sought to investigate the neurological processes underlying nodal number in equivalence formation. A simple within-subject design was used to establish two six-member equivalence classes using a MTS paradigm with a serial training structure. An EEG measure was added during the lexical decision task identical to the one employed in Experiment 7 that comprised of between baseline discrimination training and equivalence testing. 31/42 participants formed equivalence at 80% or above accuracy. The analysis of RTs indicated that priming effects were significant across all trial types, except for the 4-node trials. The largest priming effect was found on directly trained trials, the differences between certain trial types were approaching significance (i.e. DT vs. Symmetry; 1N vs. 2N). The analysis of ERPs from 8 to 13 Hz indicated that significantly greater positivity in 4-node trials compared to 1-node trials emerged around 250 to 350 ms after the target onset on middle and right hemisphere sites. Significantly greater positivity on equivalence trials compared to class inconsistent trials emerged and directly trained trials were found to be around 250 to 350 ms after the target onset on site C3. Moreover, significantly greater negativity on 4-node trials compared to 1-node trials was found around 350 to 550 ms after target onset on middle and right hemisphere sites. Significantly greater negativity in equivalence trials compared to the directly trained trials was found around 350 to 550 ms after target onset on site PZ. Differences between equivalence and class inconsistent trials were found to be approaching significance on the left, middle and right hemisphere sites. No differences were found in measures of RTs for participant who did not form equivalence.

6.6 Theoretical Issues

The aim of the current chapter was try to pinpoint the neurological processes underlying nodal distance in equivalence class formation. Despite the different nature of the experimental design, several broad findings were consistent in Experiment 7 and 8. One major consistency was the lack of the N400 in class inconsistent trials in both experiments. This finding is consistent with Haimson, et al. (2009)'s Experiment 2, in which, N400 effect only observed in participants who had received the electrophysiological testing after equivalence testing, but not before. As Experiment 7 and 8 both employed electrophysiological testing before equivalence testing, therefore the lack of N400 effect might be due to the lack of the testing contingency which provided a crucial context for the emergence of equivalence relations (Sidman, 1985). The two studies reported here not only replicated the finding of Haimson, et al. (2009)'s Experiment 2, but also extended the findings from 8 participants to 70 participants. However, this finding is contradictory with Barnes-Holmes, et al. (2005)'s Experiment 3. These authors argued that mediated priming effects can only be assessed when an equivalence test is not presented before the electrophysiological testing. Class consistent stimulus pairs were presented repeatedly during testing equivalence, therefore, it is difficult to define the priming effect observed as "direct priming" or "mediated priming". This is fatal when assuming equivalence (Sidman, 1994, 2000) or derived stimulus relations (Hayes, et al., 2001) can provide a behavioural account of semantic network formation, as only derived stimulus relations could account for the connections that have not been directly associated in a semantic network. Of course, there might be other procedure differences that account for the lack of N400 effects in Experiment 7 and 8. For example, Barnes-Holmes, et al.'s (2005) Experiment 3 used two four-member equivalence classes, Yorio, et al., (2008) used two three-member equivalence classes, both with a within subject experimental design. However, Experiment 7 used a 2x2 between-subject design with the manipulation of reinforcement and trial presentation that may have affected the formation of equivalence classes and masked the N400 effect with direct learning imposed by reinforcement and trial presentation. Hence, only directly trained trials showed the N400 (Experiment 7) rather than equivalence trials and class inconsistent trials. Experiment 8 employed a within subject design but with extended equivalence classes, in line with Haimson, et al. (2009), two six-member equivalence classes were used. The analysis of mean amplitude suggested that the N400 effect was masked by the control of nodal number in the extended equivalence classes.

Another consistent finding across the two experiments was that mean amplitudes demonstrated inflated positivity 300 ms after the target onset, which is in line with the findings from Yoiro, et al., (2008). This finding supports the postulate that the P300 is associated with source allocation, categorization, and decision making (Azizian, Freitas, Watson, & Squires, 2006). However, there were differences in the P300 reported in Experiment 7 and 8. In Experiment 7, the P300 was associated with the level of effort participants put into the task. That is, the difficulty of the task was influenced by reinforcement and trial presentation. Whereas in Experiment 8, the P300 seems associated with categorization and the level of effort participants put into the task, which is in line with the assumptions of the nodal number effect.

Both experiments confirmed the prediction that directly trained trials were the most sensitive to the effects of experimental manipulation on measures of RTs and ERPs (Experiment 7), compared to equivalence and class inconsistent trials, providing evidence for the importance of empirical training to the formation of a semantic network. Only participants who formed equivalence classes demonstrated priming effects in both experiments, the lack of priming effects in participants who did not form equivalence was in line with previous studies (Barnes-Holmes, et al., 2005 Exp2), and supported the assumption that equivalence (Sidman 1994, 2000) or derived stimulus relations (Hayes, et al., 2001) provide a behavioural account of semantic network formation. Both experiments reported major ERP changes based on experimental variables during the 8 – 13 Hz frequency range and little or no changes from 4 – 8 Hz, 13 – 22 Hz.

Additionally, a greater N400 was found on 4-node rather than 1-node trials in central, parietal and occipital lobes in Experiment 8. This can not be due to some uncontrolled event, since it was consistent with the lack of priming on 4-node trials only when measuring RTs across trial types in the same experiment. This consistency seems in line with the nodal number hypothesis, that is, as nodal number increases, stimulus control decreases, therefore, resulting in the weakest priming. This indicated a greater sense of mismatching in 4-node relations than 1-node relations.

The analysis of RTs in both experiments has produced additional interesting findings. In Experiment 7, the strength of priming was sensitive to the impact of reinforcement and trial presentation. Because the only condition that did not show a priming effect was the low reinforcement and a low number of trial presentations for passed participants. This finding provided the first empirical evidence for the assumption that reinforcement and number of trial presentations (i.e., Chapter 2) are the primary contributing factors in the formation of semantic networks. In Experiment 8, although priming effects were found across all node relations, it did not follow a linear pattern emerged in accordance with nodal number. This confirmed the speculation that the nodal number effect does not necessarily have to involve linear relations (Sidman, 1994). However, the priming effect demonstrated in 4-node relations was the only one that did not reach statistical significance when compared to other node relations (e.g. 1, 2, and 3-node), this lack of significance seems partially consistent with the assumption of nodal numbers, that is, the strength of priming is greater in stimulus pairs that have less nodes intercepted between them whereas smaller in stimulus pairs that have more nodes intercepted between them (Fields & Moss, 2007).

6.7 Suggestion for Future Study

The findings of Chapter 2 and 3 suggested that nodal number effects should account for the delayed emergence of equivalence performance when the test format has been taken into account. Therefore, future studies investigating equivalence class formation should not only consider class size, the number of comparison stimuli, the number of baseline trial presentations and the reinforcement history, but also the test format that maximizes either between class discrimination or class based discriminations. For example, the methods outlined in Experiment 5 provided a good example for assessing the relative power between the nodal and discrimination accounts using a MTS paradigm, however, a function transfer test could be added at the end of this preparation to further compare the two accounts in terms of post-equivalence performance. Additionally, increasing the number of equivalence classes which result in an increased magnitude of power in terms of the discriminative account might provide more interesting results (i.e., three five-member equivalence classes result in *four* times more presentation of stimuli B, C, and D, than the A stimulus; and 75% more than the E stimulus).

Chapter 4 provided preliminary evidence of RTs in assessing the strength of connections among equivalent stimuli. Although the calculation of mean RTs indicated significant differences in main effects and interactions, no statistically significant differences were reported in the follow-up tests. Indeed, a strong equivalence performance at the end of the experiment would be necessary to examine how robust equivalence can account for semantic network growth. Further research can address this issue by simply repeating the training phase and thus resulting in increased baseline training. Extending and balancing the number of participants who formed equivalence classes across conditions is a potential avenue for future work. Moreover, equivalence is not the only form of derived relations, there are many other derived relations (e.g. hierarchy, more-less relation, same-opposite relation, etc) that should be explored in terms of semantic processes. For example, Whelan, Cullinan, O'Donovan, and Valverde (2005) have provided evidence of same-opposite relational responding in mediated priming.

No robust N400 effect was found in Experiment 7 and 8 using EEG measures before equivalence testing, it would be worthwhile to divide participants into two groups with EEG measures either pre or post equivalence testing to assess whether the lack of N400 is a result of different procedures. The 2x2 design of Experiment 7 rendered the interpretation of the neurological data less straightforward than the within subject design employed in Experiment 8. Therefore, future experiments should consider a within subject experimental design to simplify interpretation of the findings. All experiments that involved a lexical decision task in the current thesis used short prime-target onset asynchronies (SOAs), it would be interesting to see whether long SOAs would have any effect on the strength of priming. Finally, all EEG data were average referenced in Experiment 7 and 8; it would be interesting to see whether referencing based on mastoids would produce robust N400 effects.

6.8 Concluding Comments

The current thesis aimed to provide a behavioural account of semantic network growth. It started with the systematic manipulation of training structure and reinforcement in equivalence class formation, suggesting that nodal number effects are not a result of unequal reinforcement during conditional discrimination training, rather a genuine effect expressed in terms of the test format employed. This result was further confirmed by a comparative study between the nodal number account and the discrimination account of equivalence formation, that is, the nodal number effect accounted for more equivalence class formation compared to unequal stimulus presentation during MTS training. The latter half of the thesis incorporated procedures from the semantic network and cognitive neuroscience literature by integrating cognitive and neurological measurements with behavioural principles. Three major findings emerged from the current thesis. First, the relatedness among stimuli does not appear to be a function of differential reinforcement during baseline training. Second, the relatedness among stimuli does not appear to be a function of differential discrimination during baseline training. Third, priming effects using intra-experimentally trained stimuli appear similar to priming effects reported in the cognitive literature. As a whole the current thesis added to the literature on derived stimulus relation as a behavioural account of semantic network growth.

References

Adams, B.J., Fields, L., & Verhave, T. (1993). The effects of test order on equivalence class formation. *The Psychological Record, 43*, 133-152.

Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, *32,* 177-190.

Anderson, J. R. (1976). *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Annett, J. M., & Leslie, J. C. (1995). Stimulus equivalence classes involving olfactory stimuli. *The Psychological Record*, *45*, 439-450.

Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record*, *47*, 309-320.

Augustson, E. M., & Dougher, M. J. (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *Journal of Behavior Therapy & Experimental Psychiatry*, *28*, 181-191.

Azizian, A., Freitas, A. L., Watson, T. D., & Squires, N. K. (2006). Electrophysiological correlates of categorization: P300 amplitude as index of target similarity. *Biological Psychology*, *71*, 278-288.

Barnes, D., & Hampson, P. J. (1993). Stimulus equivalence and connectionism: Implications for behavior analysis and cognitive science. *The Psychological Record*, *43*, 617-638.

Barnes, D., & Holmes, Y. (1991). Radical behaviorism, stimulus equivalence, and human cognition. *The Psychological Record, 41,* 19-31.

Barnes, D., & Keenan, M. (1993). A transfer of functions through derived arbitrary and non-arbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior*, *59*, 61-81.

Barnes-Holmes, Y., Barnes-Holmes, D., Smeets, P. M., & Luciano, C. (2004) A derived transfer of mood functions through equivalence relations. *The Psychological Record*, *54,* 95-113.

Barnes-Holmes, D., Keane, J., Barnes-Holmes, Y., & Smeets, P. M. (2000). A derived transfer of emotive functions as a means of establishing differential preferences for soft drinks. *The Psychological Record*, *50*, 493-512.

Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., Smeets, P. M., & Dymond, S. (2005). Derived stimulus relations, semantic priming, and event-related potentials: Testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior*, *84*, 417-433.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *50*, 560-585.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brian Science*, *22*, 577-660.

Beasty, A. (1987). The role of language in the emergence of equivalence relations: A developmental study. PhD thesis from University of Wales, Bangor.

Bentall, R. P., Dickins, D. W., & Fox, S. R. A. (1993). Naming and equivalence: response latencies for emergent relations. *The Quarterly Journal of Experimental Psychology*, *46B*, (May).

Bentall, R. P., Jones, R. M., & Dickins, D. W. (1998). Errors and response latencies as a function of nodal distance in 5-member equivalence classes. *Psychological Record*, *48*, 93-115.

Berger, H. (1929). On the Electroencephalogram of Man. *Journal fur Psychologie und Neurologie*, *40,* 160–179.

Bloomfield, L. (1914). *An Introduction to the Study of Language*. New York: Henry Holt and Company.

Boelens, H. (2002). Studying stimulus equivalence: Defense of the two-choice procedure. *The Psychological Record*, *52*, 305-314.

Bond, A. B., Kamil, A. C., & Balda, R. P. (2003). Social complexity and transitive inference in corvids. *Animal Behaviour*, *65,* 479-487.

Bryant, P. E., Trabassso, T., (1971). Transitive inferences and memory in young children. *Nature*, *232*, 456-458.

Brysbaert, M. (1996). Word frequency affects naming latency in Dutch with age of acquisition controlled. *European Journal of Cognitive Psychology*, *8*, 185-193.

Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive*

Psychology, 12, 65-85.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition of words is a significant variable in semantic tasks. *Acta Psychologica, 104,* 215-226.

Burgos, J. E. (2003). Laudable goals, interesting experiments, unintelligible theorizing: A critical review of Steven C. Hayes, Dermot Barnes-Holmes, and Bryan Roche's (Eds.) *Relational Frame Theory* (New York: Kluwer Academic/Plenum, 2001). *Behavior and Philosophy, 31,* 19–45.

Bush, K. M. (1993). Stimulus equivalence and cross-modal transfer. *The Psychological Record, 43,* 567-584.

Bush, K. M., Sidman, M., & de Rose, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 51,* 29-45.

Cahill, J., Barnes-Holmes, Y., Barnes-Holmes, D., Rodriguez-Valverde, M., Luciano, C., & Smeets, P. M. (2007). The derived transfer and reversal of mood functions through equivalence relations: II. *The Psychological Record, 57,* 373-389.

Cahn, B. R., & Polich, J. (2006). Meditation states and traits: EEG, ERP, and neuroimaging studies. *Psychological Bulletin. 132,* 180-211.

Carr, D., & Felce, D. (2000) Application of stimulus equivalence to language intervention for individuals with severe linguistic disabilities. *Journal of Intellectual & Developmental Disability, 25,* 181-205.

Carr, D., Wilkinson, K. M., Blackman, D., & McIlvane, W. J. (2000) Equivalence classes in individuals with minimal verbal repertoires. *Journal of the Experimental Analysis of Behavior, 74,* 101-114.

Carrigan, P. F. Jr., & Sidman, M. (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58,* 459-504.

Catania, A. C. (1998). *Learning* (4th Ed.). Upper Saddle River, NJ: Prentice-Hall.

Chance, P. (2008). *Learning and Behavior: Active Learning Edition,* (6th Ed.), Wdsworth, Cengage Learning.

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

Christopher, A. B. (1988). *Predisposition versus experiential models of compulsive* *gambling: An experimental analysis using pigeons*. Unpublished Ph.D. dissertation, West Virginia University, Morgantown.

Cofer, C. N., & Foley, J. P. (1942). Mediated generalization and the interpretation of verbal behavior: I. Prologomena. *Psychological Review, 49,* 513-540.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407-428.

Cotton, J. W. (1953). Running time as a function of amount of food deprivation. *Journal of Experimental Psychology, 46,* 188-198.

Cullinan, V. A., Barnes, D., Hampson, P. J., & Lyddy, F. (1994). A transfer of explicitly and non-explicitly trained sequence responses through equivalence relations: An experimental demonstration and connectionist model. *The Psychological Record, 44,* 559-585.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996) A neural basis for lexical retrieval, *Nature, 380,* 499-505.

D'Amato, M. R., Colombo, M., (1990). The symbolic distance effect in monkeys (*Cebus paella*). *Animal Learning & Behavior, 18,* 133-140.

Davis, S. M., & Kelly, M. H. (1997). Knowledge of English noun-verb stress difference by native and nonnative speakers. *Journal of Memory and Language, 36,* 445-460.

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition, *Psychophysiology, 41*, 60–74.

Deacon, T. (1997). *The Symbolic Species. The Co-evolution of Language and the Human Brain.* London: Penguin Books.

De Grandpre, R. J., Bickel, W. K., & Higgins, S. T. (1992). Emergent equivance relations between interoceptive (drug) and exteroceptive (visual) stimuli. *Journal of the Experimental Analysis of Behavior*, *58*, 9-18.

De Rose, J. C., McIlvane, W. J., Dube, W. V., Galpin, V. C., & Stoddard, L. T. (1988). Emergent simple discrimination established by indirect relation to differential consequences. *Journal of the Experimental Analysis of Behavior, 50*, 1-20.

Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986) Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior, 46,* 243-257.

Dickins, D. W. (2005). On aims and methods in the neuroimaging of derived relations. *Journal of the Experimental Analysis of Behavior, 84,* 453-483.

Dickins, D. W., Bentall, R. P., & Smith, A.B. (1993). The role of individual stimulus names in the emergence of equivalence relations: The effects of interpolated paired-associates training of discordant associations between names. *The Psychological Record, 43,* 713-724.

Donchin, E., & Coles, M. G. H. (1988) Is the P300 component a manifestation of cognitive updating? *The Behavioral and Brain Sciences, 11,* 357–427.

Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). Transfer of respondent elicitation and avoidance evocation through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 62,* 331-351.

Dube, W. V., Green, G., & Serna, R. W. (1993). Auditory successive conditional discrimination and auditory stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 59,* 103-114.

Dube, W. V., & McIlvane, W. J. (1989). Adapting a microcomputer for behavioral evaluation of mentally retarded individuals. In J. A. Mulick & R. F. Antonack (Eds.), *Transitions in mental retardation.* Norwood, NJ: Ablex.

Dube, W. V., McIlvane, W. J., Mackay, H. A., & Stoddard, L. T. (1987). Stimulus class membership established via stimulus-reinforcer relations. *Journal of the Experimental Analysis of Behavior, 47,* 159-175.

Ducharme, D. E., & Holborn, S. W. (1997). Programmed generalization of social skills in preschool children with hearing impairments. *Journal of Applied Behavior Analysis, 30(4),* 639-651.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. LeJeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Dymond S., & Barnes, D. (1995). A transformation of self-discrimination response functions in accordance with the arbitrarily applicable relations of sameness, more than, and less than. *Journal of the Experimental Analysis of Behavior, 64,* 163-184.

Dymond, S., Gomez-Martin, S., & Barnes, D. (1996). Multi-modal conditional discrimination in rats: Some preliminary findings. *The Irish Journal of Psychology, 17,* 269-281.

Dymond, S. & Rehfeldt, R. (2000). Understanding complex behaviour: The transformation of stimulus functions. *The Behavior Analyst, 23*, 239-254.

Dymond, S. & Rehfeldt, R.A. (2001). Supplemental measures of derived stimulus relations. *The Experimental Analysis of Human Behavior Bulletin, 19*, 6-10.

Dymond, S., Roche, B., Forsyth J. P., Whelan R., & Rhoden J. (2007). Transformation of avoidance response functions in accordance with same and opposite relational frames. *Journal of the Experimental Analysis of Behavior, 88*, 249-262.

Ellis, A .W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Leaning, Memory, and Cognition, 24*, 515-523.

Escobar, R. & Bruner, C. A. (2007). Response induction during the acquisition and maintenance of lever pressing with delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 88*, 29-49.

Fields, L., Adams, B., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53,* 345-358.

Fields, L., Adams, B.J., Verhave, T., Newman, S. (1993). Are stimuli in equivalence classes equally related to each other? *The Psychological Record, 43*, 85-105.

Fields, L., Landon-Jimenez, D. V., Buffington, D. M., Adams, B. J. (1995). Maintained nodal-distance effects in equivalence classes. *Journal of the Experimental Analysis of Behavior, 64,* 129-145.

Fields, L., & Moss, P. (2007) Stimulus Relatedness in Equivalence Classes: Interaction of Nodality and Contingency. *European Journal of Behavior Analysis, 8,* 141 – 159.

Fields, L., Reeve, K. F., Rosen, D., Varelas, A., Adams, B. J., Belanich, J., & Hobbie, S. A. (1997). Using the simultaneous protocol to study equivalence class formation: The facilitating effects of nodal number and size of previously established equivalence classes. J*ournal of the Experimental Analysis of Behavior, 67*, 367–389.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 48,* 317-332.

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of the Experimental Analysis* *of Behavior, 42*, 143-157.

Fields, L., & Watanabe-Rose, M. (2008). Nodal Structure and the Partitioning of Equivalence Classes. *Journal of the Experimental Analysis of Behavior, 89*, 359- 381.

Gerband, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Leaning, Memory, and Cognition, 24*, 267-283.

Gerhand, S., & Barry, C. (1999). Age-of-acquisition and frequency effects in speeded word naming. *Cognition, 73*, 27-36.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica, 115*, 43-67.

Girden, E. R. (1992). *ANOVA: Repeated Measures*. Sage University Papers Series on Quantitative Applications in the Social Sciences, *84*. Thousand Oaks, CA: Sage.

Goyos, C. (2000) Equivalence class formation via common reinforcers among preschool children. *The Psychological Record, 50,* 629-654.

Gray, H. M., Ambady, N., Lowenthal, W. T., & Deldin, P. (2004). P300 as an index of attention to self-relevant stimuli, *Journal of Experimental Social Psychology, 40,* 216-224.

Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation, 95*, 260–270.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? *Psychological*

Bulletin, 83, 314-320.

Grover, D. E., Horton, D. L., & Cunningham, M., Jr. (1967). Mediated facilitation and interference in a four-stage paradigm. *Journal of Verbal Learning and Verbal Behavior, 6*, 42-46.

Gutiérrez,-Martínez, O., Luciano-Soriano, M.C., & Valdivia-Salas, S. (2005). Change of self-efficacy verbalizations and derivation of functions. *Psicothema, 17,* 614-619.

Guttman, N. (1963). Laws of behavior and facts of perception. In S. Koch (Ed.), *Psychology: A study of a science* (Vol.5). Now York: McGraw-Hill.

Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization.

Journal of Experimental Psychology, *51*, 79-88.

Haimson, B., Wilkinson, K. M., Rosenquist, C., Ouimet, C., & McIlvane, W. J. (2009). Electrophysiological correlates of stimulus equivalence processes. *Journal of the Experimental Analysis of Behavior*. *92*, 245-256.

Hammond, L. J. (1980). The effect of contingency upon the appetitive conditioning of free-operant behaviour, *Journal of the Experimental Analysis of Behavior*, *34(3)*, 297-304.

Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, *58*, 321-334.

Hayes, S. C. (Ed.). (1989). *Rule-governed behavior: Cognition, contingencies, and instructional control*. New York: Plenum.

Hayes, S. C. (1994). Relational frame theory: A functional approach to verbal events. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition.* Reno, NV: Context Press.

Hayes, S. C., & Barnes-Holmes, D. (2004). Relational operants: Processes and implications: A response to Palmer's review of relational frame theory. *Journal of the Experimental Analysis of Behavior*, *82*, 213-224.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition*, New York: Plenum.

Hayes, S. C., & Bisset, R. T. (1998). Derived stimulus relations produce mediated and episodic priming. *The Psychological Record*, *48*, 617-630.

Hayes, S. C., Brownstein, A. J., Devany, J. M., Kohlenberg, B. S., & Shelby, J. (1987). Stimulus equivalence and the symbolic control of behavior. *Mexican Journal of Behavior Analysis*, *13*, 361-374.

Hayes, S. C., Fox, E., Gifford, E. V., Wilson, K. G., Barnes-Holmes, D., & Healy, O. (2001). Derived relational responding as learned behavior. In S. C. Hayes, D. Barnes-Holmes, & B. Roche, (Eds.) *Relational frame theory: A post-Skinnerian account of human language and cognition*, New York: Plenum.

Hayes, S. C., & Hayes, L. J. (Ed.). (1992). *Understanding verbal relations*. Reno, NV: Context Press.

Hayes, L. J., Thompson, S., & Hayes, S. C. (1989). Stimulus equivalence and rule following. *Journal of the Experimental Analysis of Behavior*, *52*, 275-291.

Hayes, L. J., Tilley, K. J., & Hayes, S. C. (1988). Extending equivalence class membership to gustatory stimuli. *The Psychological Record, 38,* 473-482.

Heinze, H., Muente, T., & Kutas, M. (1998). Context effects in a category verification task as assessed by event-related brain potential (ERP) measures. *Biological Psychology*, *47*, 121–135.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13,* 243-266.

Hinojosa, J. A., Martín-Loeches, M., & Rubia, F. J. (2001). Event-related potentials and semantics: an overview and an integrative proposal. *Brain and language*, *78,* 128-39.

Hogan, D. E., & Zentall, T. R. (1977). Backward associations in the pigeon. *American Journal of Psychology*, *90*, 3-15.

Holmes, P. W. (1979). Transfer of matching performance in pigeons. *Journal of the Experimental Analysis of Behavior*, *31*, 103-114.

Holth, P., & Arntzen, E. (2000). Reaction times and the emergence of class consistent responding: A case for precurrent responding? *The Psychological Record, 50,* 305-337.

Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, *65*, 185-241.

Hovland, C. I. (1937). The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *Journal of General Pscyhology*, *17,* 125-148.

Hugdahl, K. (1995/2001). *Psychophysiology: The Mind-body Perspective.* Cambridge, MA: Harvard University Press.

Imam, A. A. (2001). Speed contingencies, number of stimulus presentations, and the nodal number effect in equivalence class formation. *Journal of the Experimental Analysis of Behavior, 76*, 265-288.

Imam, A. A. (2003). Assessing transfer of response speed and nodal number via conditional discriminations. *Experimental Analysis of Human Behavior Bulletin, 21*, 1-7.

Imam, A. A. (2006). Experimental control of nodal number via equal presentations of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of the Experimental Analysis of Behavior, 85*, 107-24.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for Multiple resources in dual-task performance. *Psychophysiology, 17,* 259-273.

Izura, C. & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicologica, 23*, 245-281.

James, C. T., & Hakes, D. T. (1965). Mediated transfer in a four-stage, stimulus equivalence paradigm. *Journal of Verbal Learning and Verbal Behavior, 45,* 297-304.

Jenkins, J. J. (1963). Mediated associations: paradigms and situations. In C. N. Cofer, and B. S Musgrave (Eds.), Verbal behavior and learning: Problems and processes. New York: McGraw-Hill.

Jenkins, J. J. (1965). Mediation theory and grammatical behavior. In S. Rosenberg (Ed.), *Directions in psycholinguistics.* New York: Macmillan.

Jenkins, J. J., & Palermo, D. S. (1964). Mediation processes and the acquisition of linguistic structure. In U. Bellugi & R. Brown (Eds.), *The acquisition of language. Monographs of the Society for Research in Child Development, 29,* 141-169.

Johnson, C., & Sidman, M. (1993). Conditional discrimination and equivalence relations: Control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 59,* 333-347.

Johnson, R., Jr. (1984). P300: A model of the variables controlling its amplitude. *Annals of the New York Academy of Sciences, 425,* 223-229.

Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology, 23,* 367-384.

Johnson, R., Jr. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis, *Advances in Psychophysiology, 3,* 69–137.

Kennedy, C. L. (1991). Equivalence class formation influenced by the number of nodes separating stimuli. *Behavioural Processes, 24,* 219-245.

Kennedy, C. H., Itkonen, T., & Lindquist, K. (1994). Nodality effects during equivalence class formation: An extension to sight-word reading and concept development. *Journal of Applied Behavior Analysis, 27,* 673–683.

Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y. H., Emre, M., & Demiralp, T. (2006). Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain Research, 1104,* 114–128.

Kisley, M. A., & Cornwell, Z. M. (2006). Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clinical Neurophysiology, 117,* 2549-2563.

Kucera, H., & Francis, W. N. (1967). *Computational Analysis of Present-day American English.* Providence, RI: Brown University Press.

Kutas, M. & Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203-205.

Kutas, M. & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition, 11,* 539-550.

Kutas, M. & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.

Lacey, J. I., Smith, R. L., & Green, A. (1955). Use of conditioned autonomic responses in the study of anxiety. *Psychosomatic Medicine, 17,* 208-217.

Lattal, K. A. (1995). Contingency and behavior analysis, *The Behavior Analyst, 18,* 209-224.

Lazar, R. (1977). Extending sequence-class membership with matching to sample. *Journal of the Experimental Analysis of Behavior, 27,* 381-392.

Lazar, R. M., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior, 41,* 251-266.

Leslie, J. C., Tierney, K. J., Robinson, P., Keenan, M., Watt, A., & Barnes, D. (1993). Differences between clinically anxious and non-anxious subjects in a stimulus equivalence training task involving threat words. *The Psychological Record, 43,* 153-161.

Lewis, M. B. (1999). Age of acquisition in face categorisation: Is there an instance-based account? *Cognition, 71,* 23-39.

Lipkens, G., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology, 56,* 201-239.

Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5,* 161-171.

Lowe, C. F. (1986). *The role of verbal behavior in the emergence of equivalence* *relations.* Paper presented at the meeting of the Association for Behavior Analysis, Milwaukee, WI.

Lowe, C. F., & Beasty, A. (1987). Language and the emergence of equivalence relations: A developmental study. *Bulletin of the British Psychological Society, 40,* A42.

Luck, S. J. (2005). *An Introduction to the Event-related Potential Technique*, MIT Press.

Ludvig, E. A., Conover, K., & Shizgal, P. (2007). The effects of reinforcer magnitude on timing in rats. *Journal of the Experimental Analysis of Behavior*, *87*, 201-218.

Manabe, K., Kawashima, T., & Staddon, J. E. R. (1995). Differential vocalization in budgerigars: Towards an experimental analysis of naming. *Journal of the Experimental Analysis of Behavior, 63*, 111-126.

Martin, A., Ungerleider, L.G., Haxby, J.V., (2000) Category-specificity and the brain: the sensory-motor model of semantic representations of objects. *The cognitive neurosciences*, (2nd Ed) (Gazzaniga, M. S., Ed.), Cambridge, MA: MIT Press.

Massaro, D. W., Venezky, R. L., & Taylor, G. A. (1979). Orthographic regularity, positional frequency, and visual processing of letter strings. *Journal of Experimental Psychology: General, 108*, 107-124.

McClelland, J. L., & Rumelhart, D. E. (1988). *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. Cambridge, MA: MIT Press.

McDonagh, E. C., McIlvane, W. J., & Stoddard, L. T. (1984). Teaching coin equivalences via matching to sample. *Applied Research in Mental Retardation, 5,* 177–197.

McGill, P. (1999). Establishing operations: Implications for the assessment, treatment, and prevention of problem behavior. *Journal of Applied Behavior Analysis, 32,* 393- 418.

McIlvane, W. J., Dube, W. V., Kledaras, J. B., de Rose, J. C., & Stoddard, L. T. (1992). Stimulus-reinforcer relations and conditional discrimination. In S. C.

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37,* 149–155.

Michael, J. (1988). Establishing operations and the mand. *The Analysis of Verbal Behavior, 2,* 19–21.

Michael, J. (2000). Implications and refinements of the establishing operation concept, *Journal of Applied Behavior Analysis, 33,* 401–410.

Minster, S. T., Jones, M., Elliffe, D., & Muthukumaraswamy, S. D. (2006). Stimulus equivalence: Testing Sidman's (2000) Theory. *Journal of the Experimental Analysis of Behavior, 85,* 371-391.

Miyashita, Y., Nakajima, S., & Imada, H. (2000). Differential outcome effect in the horse. *Journal of the Experimental Analysis of Behavior, 74,* 245-253.

Moore, V., & Valentine, T. (1998). Naming faces: The effect of AoA on speed and accuracy of naming famous faces. *The Quarterly Journal of Psychology, 51,* 458-513.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Leaning, Memory, and Cognition, 21,* 116-133.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91,* 167-180.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processing in reading: Visual word recognition.* Hillsdale, NJ: Erlbaum.

Oldfiled, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia, *9,* 97-113.

Palmer, D. C. (2004a). Data in search of a principle: A review of S. C. Hayes, D. Barnes-Holmes, and B. Roche (Eds.), *Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition. Journal of the Experimental Analysis of Behavior, 81,* 189–204.

Palmer, D. C. (2004b). Generic response classes and relational frame theory: response to Hayes and Barnes-Holmes. *Journal of the Experimental Analysis of Behavior, 82,* 225-234.

Pavlov, I. P. (1927). *Conditioned Reflexes* (G. V. Anrep, Ed. and Trans.). London: Oxford University Press.

Peterson, G. B., & Trapold, M. A. (1980). Effects of altering outcome expectancies on pigeons' delayed conditional discrimination performance. *Learning and Motivation, 11,* 267-288.

Pfurtscheller, G., & Lopes, da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology, 110,* 1842–1857.

Pilgrim, C., & Galizio, M. (1995). Reversal of baseline relations and stimulus equivalence: I. Adults. *Journal of the Experimental Analysis of Behavior*, *63*, 225–238.

Powers, R. B., Cheney, C. D., & Agostino, N. R. (1970) Errorless training of a visual discrimination in preschool children. *Psychological Record*, *20*, 45-50.

Purtle, R. B. (1973). Peak shift: A review. *Psychological Bulletin*, 80, 408-421.

Purves, D., Brannon, E. M., Cabeza, R., Huettel, S., LaBar, K. S., Platt, M. L., & Woldorff, M. G. (2008). *Principles of Cognitive Neuroscience*, Sinauer Associates, Inc.

Randall, T., & Remington, B. (1999). Equivalence relations between visual stimuli: The functional role of naming. *Journal of the Experimental Analysis of Behavior*, *71*, 395-416.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*, 510-532.

Razran, G. (1939). A quantitative study of meaning by a conditioned salivary technique (semantic conditioning). *Science*, *90*, 89-90.

Rehfeldt, R. A., & Dymond, S. (2005). The Effects of Test Order and Nodal Distance on the Emergence and Stability of Derived Discriminative Stimulus Functions. *The Psychological Record, 55,* 179-196.

Reynolds, W. F., & Pavlik, W. B. (1960). Running speed as a function of deprivation period and reward magnitude. *Journal of Comparative and Physiological Psychology*, *53*, 615-618.

Roche, B., & Barnes, D. (1997). A transformation of respondently conditioned stimulus function in accordance with arbitrarily applicable relations. *Journal of the Experimental Analysis of Behavior*, *67*, 275-301.

Roche, B., Barnes-Holmes D., Smeets P. M., Barnes-Holmes Y., & McGeady S. (2000). *Contextual control over the derived transformation of discriminative and sexual arousal functions. The Psychological Record, 50,* 267-291.

Rodewald, H. K. (1974). Symbolic matching-to-sample by pigeons. *Psychological Reports*, *34*, 987-990.

Saunders, K. J., Williams, D. C., & Spradlin, J. E. (1996). Derived stimulus control: Are there differences among procedures and processes? In T. R. Zentall & P.

M. Smeets (Eds.), *Stimulus class formation in humans and animals.* Amsterdam, NL: Elsevier.

Saunders, R. R., Chaney, L., Marquis, J. G. (2005). Equivalence class establishment with two-, three-, and four-choice matching to sample by senior citizens. *The Psychological Record, 55*, 539-559.

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of the Experimental Analysis of Behavior, 71*, 195-214.

Saunders, R. R., & Green, G. (1992) The non-equivalence of behavioral and mathematical equivalence. *Journal of the Experimental Analysis of Behavior, 57,* 227–241.

Saunders, R. R., & Green, G. (1996). Naming is not (necessary for) stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65*, 312–314.

Saunders, R. R., & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of the Experimental Analysis of Behavior, 72*, 117-137.

Semon, R. (1921). *The Mneme*, London: George Allen & Unwin.

Schlinger, H. D., & Blakely, E. (1994). The effects of delayed reinforcement and a response-produced auditory stimulus on the acquisition of operant behavior in rats, *The Psychological Record, 44*, 391-409.

Schneider, J. W. (1973). Reinforcer effectiveness a function of reinforcer rate and magnitude: A comparison of concurrent performance. *Journal of the Experimental Analysis of Behavior, 20*, 461-471.

Schusterman, R. J., & Kastak, D. (1993) A california sea lion (*Zalophus Californianus*) is capable of forming equivalence relations, *The Psychological Record, 43,* 823-839.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*, 5-13.

Sidman, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Sidman, M. (1992). Equivalence relations: Some basic considerations. In S. C. Hayes & L. J. Hayes (Eds.), *Understanding verbal relations*. Reno, NV: Context Press.

Sidman, M. (1994). *Stimulus equivalence: A research story*. Boston: Authors Cooperative.

Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior, 74*, 127-146.

Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalence in severe retardation. *American Journal of Mental Deficiency, 77,* 515-523.

Sidman, M., Cresson, O., Jr., & Willson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior*, *22*, 261-273.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior*, *43,* 21-42.

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior*, *37*, 23-44.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37,* 5-22.

Sidman, M., & Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities*, *6,* 1-19.

Siemann, M., Delius, J. D., & Wright, A. A. (1996). Transitive responding in pigeons: influences of stimulus frequency and reinforcement history. Behavioural Processes, *37,* 185-195.

Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*, New York: Appleton-Century-Crofts.

Skinner, B. F. (1953). *Science and Human Behavior.* Pearson Education, Inc.

Skinner, B. F. (1957). *Verbal Behavior*, New York: Appleton-Century-Crofts.

Smeets, P. M., & Barnes-Holmes, D. (2003). Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer. *Journal of Economic Psychology*, *24,* 603-618.

Smeets, P. M., Barnes-Holmes, D., & Cullinan, V. (2000) Establishing equivalence classes with match-to-sample format and simultaneous-discrimination format conditional discrimination tasks. *The Psychological Record, 50*, 721-744.

Smeets, P. M., Roche, B., & Barnes-Holmes, D. (1997) Functional equivalence in children: derived stimulus-response and stimulus-stimulus relations. *Journal of Experimental Child Psychology, 66,* 1-17.

Smith, R. G., & Iwata, B. A. (1997). Antecedent influences on behaviour disorders, *Journal of Applied Behavior Analysis, 30*, 343–375.

Smyth, S., Barnes-Holmes, D., & Forsyth, J. (2006). A derived transfer of simple discrimination and self-report arousal functions in spider fearful and nonspider fearful participants. *Journal of the Experimental Analysis of Behavior, 85*, 223–246.

Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, *43,* 427-449.

Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review. 44,* 430-444.

Spence, K. W. (1960). *Behavior theory and learning*. Englewood Cliffs, NJ: Prentice-Hall.

Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65,* 643-659.

Spradlin, J. E., & Saunders, R. R. (1986). The development of stimulus classes using match-to-sample procedures: Sample classification vs. comparison classification. *Analysis and Intervention in Developmental Disabilities, 6,* 41–58.

Steyvers, M., & Tenenbaum, J. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29,* 41-78.

Stromer, R., Mackay, H. A., & Remington, B. (1996). Naming, the formation of stimulus classes, and applied behavior analysis. *Journal of Applied Behavior Analysis*, *29,* 409-431.

Stromer, R., & Osbourne, J. G. (1982). Control of adolescents' arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior*, *37,* 329-348.

Tabullo, A., Yorio, A., Perez-Leguizamon, P., & Segura, E. (2008). An ERP study of category leaning. *Clinical Neurophysiology*, *119*, e155.

Taylor, K.I. & Regard, M. (2003). Language in the right cerebral hemisphere: Contributions from reading studies. *News in Physiological Sciences, 18*, 257-61.

Terrace, H. S. (1963a). Discrimination learning with and without "errors". *Journal of the Experimental Analysis of Behavior, 6,* 1-27.

Terrace, H. S. (1963b). Errorless transfer of a discrimination across two continua. *Journal of the Experimental Analysis of Behavior, 6,* 223-232.

Terrace, H. S. (1964). Wavelength generalization after discrimination learning with and without errors. *Science, 144,* 78-80.

Terrace, H. S. (1972). By-products of discrimination learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 5). New York: Academic Press.

Thomas, D. R., Mood, K., Morrison, S., & Wiertelak, E. (1991). Peak shift revisited: A test of alternative interpretations. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 130-140.

Todorov, J. C., Hanna, E. S., & Bittencourt de Sa, M. C. N. (1984). Frequency versus magnitude of reinforcement: New data with a different procedure. *Journal of the Experimental Analysis of Behavior, 4*, 157-167.

Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation, 1,* 129-140.

Treichler, F. R. & Tilburg, D. V. (1996) Concurrent conditional discrimination tests of transitive inference by Macaque Monkeys: List Linking. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 105-117.

Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Memory and Cognition, 26,* 1282-1291.

Underwood, B. J. (1949). *Experimental Psychology*. New York: Appleton-Century-Crofts.

Urcuioli, P. (2001). Categorization & acquired equivalence. In R. G. Cook (Ed.), *Avian visual cognition* [On-line]. Available: www.pigeon.psy.tufts.edu/avc/urcuioli/

Vasconcelos, M. (2008). Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes, 78*, 313-334.

Ward-Robinson, J. & Hall, G. (1999), The role of mediated conditioning in acquired equivalence. *The Quarterly Journal of Experimental Psychology, 52B,* 335-350.

Watson, J. B., & Rayner, R. (1920). Conditional emotional reactions. *Journal of Experimental Psychology, 3,* 1-4.

Whelan, R. (2008). Effective Analysis of Reaction Time Data. *The Psychological Record, 58,* 475-482.

Whelan, R., Cullinan, V., O'Donovan, A., & Valverde, M. R. (2005). Derived same and opposite relations produce association and mediated priming. *International Journal of Psychology and Psychological Therapy, 5,* 247-264.

Whelan, R., Barnes-Holmes, D., & Dymond, S. (2006). The transformation of consequential functions in accordance with the relational frames of more-than and less-than. *Journal of the Experimental Analysis of Behavior, 86,* 317-335.

White, H. (1986) Semantic priming of nonwords in lexical decision, *The American Journal of Psychology, 99,* 479-485.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50,* 125-144.

Wynne, C. D. L. (1997). Pigeon transitive inference: Tests of simple accounts of a complex performance. *Behavioural Processes, 39,* 95-112.

Yorio, A., Tabullo, A., Wainselboim, A., Barttfeld, P., & Segura, E. (2008). Event-related potential correlates of perceptual and functional categories: Comparison between stimuli matching by identity and equivalence. *Neuroscience Letters, 443,* 113-118.

Zentall, T. R. (1998). Symbolic representation in animals: Emergent stimulus relations in conditional discrimination learning. *Animal Learning & Behavior, 26,* 363-377.

Zentall, T. R., & Singer, R. A. (2007). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than a less aversive event. *Journal of the Experimental Analysis of Behavior, 88,* 131-149.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47,* 1-29.

Appendices

Appendix 1 Edinburgh Handedness Inventory (Oldfield, 1971).

Appendix 1

Subject N°...........   Date............

Handedness Inventory

For each of the ten activities below, please answer:

a) which hand you prefer for that activity, and
b) b) whether you ever use the other hand for the activity

	Activity	Which hand do you prefer to use?	Do you ever use the other hand?
1	Writing		
2	Drawing		
3	Throwing		
4	Using scissors		
5	Using a toothbrush		
6	Using a knife (without fork)		
7	Using a spoon		
8	Using a broom (upper hand)		
9	Striking a match		
10	Opening a box (lid)		