USING THE CONCEPT OF RULE - GOVERNED BEHAVIOUR TO INTEGRATE THE COGNITIVE AND BEHAVIOURAL THERAPIES: A THEORETICAL ANALYSIS

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Arts in Psychology

Massey University

David Leonard Secker

1999

ABSTRACT

The concept of *rule-governed behaviour* (RGB) was introduced by B.F Skinner (1969) to allow complex verbal behaviour to be amenable to the same types of contingency analysis used in most other areas of applied behaviourism. Inherent in the concept of RGB is the notion that people formulate and follow rules created by themselves and others and that this constitutes a distinct class of operant functioning. As the process of cognitive therapy is primarily undertaken in a verbal fashion, the possibility of employing the concept of RGB to redefine aspects of cognitive therapy from an operant perspective has been considered by several researchers (e.g., Zettle & Hayes, 1982; Poppen, 1989). This form of paradigmatic integration, involving the transplantation of one set of therapeutic techniques into the theoretical body of another epistemological framework, can be termed *assimilative* (Lazarus & Messer, 1991). The present essay clarifies the aims and content of such an integration as it relates to the concept of RGB and cognitive-behavioural rapprochement, and offers several theoretical advancements in this direction. Errors in rule-following and rule-formulation are discussed in terms of the role they play in cognitive assessment, and the cognitive mechanisms involved in therapeutic change are also analysed in terms of RGB. It is also shown how RGB can be conceptualised as a reciprocally-determined system of responding, similar to that espoused by the cognitive theorist Albert Bandura (1977a). Finally, some of the problems associated with the concept of RGB and psychotherapy integration are reviewed in relation to the present analysis.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

PREFACE AND ACKNOWLEDGEMENTS

It is rare at Masters level to submit a purely theoretical dissertation. As such, the present essay bears little resemblance to the experimental projects typical of those conducted at an early postgraduate level. Nevertheless, it is at this academic level which one can most commonly find advanced courses entirely devoted to the practice of "cognitive-behavioural therapy", but with only a fraction of the curriculum reserved for debate on the core differences between the two paradigms that constitute this therapeutic intervention. A personal dissatisfaction with this omission enthused me to conduct a critical analysis of how therapists might benefit from using the concept of *rule-governed behaviour* to analyse and understand these differences. This thesis has had several minor revisions since an earlier submission.

Anyone who does not understand B.F. Skinner's theory of rule-governed behaviour and the theoretical underpinnings of the school from which it is born is unlikely to respect the nature of the approach taken here. Because of this, a considerable amount of space has been allocated to a description of the philosophical stance taken by radical behaviourists. Part One is primarily introductory in nature and is intended to solidify and organise the conceptual tools which have been used to conduct this analysis. In contrast to the behavioural literature, Part Two outlines the fundamental tenets of four of the most influential cognitive therapies and the cognitive paradigm in general. At this point, a method by which we might usefully place the process of interpreting cognitive therapy in terms of RGB within the broader field of psychotherapy integration is offered.

Thus, Parts One and Two aim to construct a grounding for the final section, Part Three, which offers an in-depth analysis of how many rule-based clinical techniques might be used as a way for behaviour therapy to "tap-in" to those areas traditionally of concern to cognitive therapists. The final subsections of Part Three contain an evaluation of the more complex empirical and theoretical difficulties facing the concept of RGB as an integrative tool and their relation to the theoretical account given here. For the reader requiring a preparatory introduction to the field of radical behaviourism, I recommend Baum's (1994) "Understanding Behaviourism: Science, Behaviour, and Culture", and Skinner's (1969) "Contingencies of Reinforcement: A Theoretical Analysis".

I would like to take the opportunity to briefly those who have contributed to the completion of this thesis. I would like to thank Dr Kevin Ronan for his excellent supervision and commentary, and more generally, for his reassuring faith in my potential over the last three years. I would also like to express my gratitude to Dr Alan Winton for his conversation and guidance, and for sharing his useful ideas and insight with me. I also give my heartfelt thanks to my parents, Rose and Harry, for their enduring support throughout the writing of this thesis and over the course of my tertiary education.

PART ONE

THE CONCEPT OF RULE-GOVERNED BEHAVIOUR

The primary purpose of Part One is to fully outline B.F. Skinner's (1969) theory of *rule-governed behaviour* (RGB) and explain how several other researchers have extended and empirically verified his theory. The section begins with some brief comments on the philosophy of radical behaviourism and then directly moves into the areas of verbal behaviour, problem-solving, and RGB. After a description of the functional units of RGB offered by Zettle and Hayes (1982), there follows a review of some of the experimental research on RGB and an account of the relationship between the concept of RGB and the *law of effect*. Several criticisms of applied behaviour analysis are then raised, and it is argued that the emergence of the cognitive paradigm can in some ways be seen as a response to these difficulties. Finally, several reasons for using the concept of RGB as an integrative theme for the cognitive and behavioural therapies are presented.

1.1 Radical Behaviourism and Operant Behaviour

As a field of psychology, behaviourism is driven by the doctrine that organismic activity can be explained, predicted, and controlled through an adherence to traditional forms of scientific inquiry. The roots of behaviourism can be traced back to the late 19th century and the emergence of two fields in the natural sciences: *objective* psychology, which first emphasised that behaviour could be quantified, and *comparative* psychology, which embraced the idea of a continuity in behavioural traits across the species (Baum, 1994). John B. Watson (1879-1958), the founder of behaviourism, combined elements of objective and comparative psychology to create research methods which could be used to measure and identify these universal behavioural patterns, sometimes referred to as *Watsonian behaviourism*.

Watson sought the discovery of stimulus-response relations that could be described in terms closely resembling those found in physiology. Watson's approach was revolutionary in that he contended that the use of concepts requiring an introspective approach, such as those referred to by the terms "mind" and "consciousness", are detrimental to a psychology wishing to establish itself as a valid natural science. Furthermore, Watson's break from the psychoanalytic tradition paved the way for

empiricist's to form an extreme philosophical variant of his own psychology, commonly referred to as the *methodological* approach. The essential feature of *methodological behaviourism* is the idea that only publicly observable behaviour ought to be entered into any scientific analysis of behaviour. For example, thoughts, feelings and emotions that can not be both directly measured and verified with interobserver agreement are inadmissible as scientific data under the methodological regime.

Methodological behaviourism can be contrasted with the philosophy of *radical behaviourism*, a separate school founded and developed by the late B.F. Skinner (1904-1990). Radical behaviourism argues that thoughts, feelings, and emotions should be studied as *private* events, which are in no special way different from *public* events, except that they can only be observed by an audience of *one* (Skinner, 1945). Private events may be harder to quantify or "reveal" of course, but they still contribute to the broad universe of possible behaviours, and thus carry the potential to be both quantified and studied in relation to controllable variables. The reason that radical behaviourists endorse this position is partly a pragmatic one, because it is felt that an analysis of private behaviour is essential if it will afford a greater understanding of human behaviour (Baum, 1994).

Radical behaviourism does, however, oppose *mentalism.* Mentalist explanations are those which refer to non-physical entities as having a direct influence on behaviour, often giving the impression of a complete causal account but essentially offering no more than an unverified assumption (e.g., psychoanalytic concepts such as the "ego" have no direct connection with any observable phenomenon, and are thus considered mentalistic). Private events may influence behaviour, but they only do so by entering into a sequence of public-private interactions, and thus all causes of behaviour must ultimately lie in the environment[1] (Skinner, 1969). Ontologically, the philosophy of Skinner's position relies on *determinism*. Determinism is the contention that all events are the direct result of other physical events, therefore implying that true randomness is theoretically impossible (Chaplin, 1985). In alliance with the deterministic position, radical behaviourists argue

[1] Equally, however, the environment is not bounded by the skin. *Inner* or private environmental events can and do control behaviour (Skinner, 1974). Thus, the distinction between public and private events is only an empirical one.

that because mentalistic explanations rely on dimensionally fictional entities as having causal properties, they seriously detract attention from many important variables of psychological interest.

Any introduction to a behaviour-analytic thesis could be rendered incomplete without an explanation of the process of *conditioning*. Conditioning is the establishment of a relationship between stimulus and response formed through a history of association. There are two major types of conditioning: *respondent* and *operant*. Respondent or *classical* conditioning is linked to the *unconditioned response*. Unconditioned responses are those behaviours which an organism has an innate capacity to exhibit, such as blinking from a sudden movement in close proximity to the face (a *reflex*), or an animal's enactment of a mating dance (a complex series of reflexes, or a *fixed-action pattern*). Unconditional responses are thought to be biologically programmed and require no prior lifetime experience with the *unconditional* stimuli that elicit such responding. All species acquire a range of these automatic reactions through natural selection. Reflexes and fixed-action patterns which have favoured some organisms over others have been genetically passed down through generations with stronger variations of that trait. Evidence for this is supported by the observation that most unconditional responses serve some sort of protective, homeostatic, or biological function (e.g., sneezing serves to remove irritants from the nasal passage). Ivan Pavlov (1849-1936) discovered that when a neutral stimulus was repeatedly paired with an unconditional stimulus (i.e., conditioned with), the neutral stimulus would eventually come to elicit a similar response to that elicited by the unconditional stimulus on its own, which he termed the *conditioned response*. Pavlov demonstrated this phenomena by showing that the repeated presentation of meat to a dog (an unconditional stimulus for the response of salivating), paired with the presence of a light (a neutral stimulus), would eventually cause the light to elicit salivating on its own (a conditioned response).

However, only a small proportion of behaviour appears to be the result of classical conditioning, and instead tends to be influenced by its *consequences*. Operant or *instrumental* conditioning establishes an increased response probability through a consequential outcome. That is, when a response occurs and is followed by a reinforcing stimulus (a reinforcer), the probability that the response will occur again becomes higher

(Skinner, 1969). A *reinforcer* is defined as any agent which strengthens the behaviour it follows (e.g., increases frequency, duration, etc). Thus, the definition is somewhat circular. Reinforcement may involve the introduction of an establishing agent following a response (positive reinforcement) or the removal of an aversive agent following a response (negative reinforcement). Conversely, *punishment* is defined as any consequential event which weakens the behaviour it follows. Like reinforcement, the response-dependent delivery of punishment can be both positive (the addition of an aversive stimulus) or negative (the removal of a reinforcing stimulus).

Usually, operant behaviour is analysed as a three-term relation, or *contingency*: (a) an antecedent stimulus or context, (b) a response, and (c) a reinforcing consequence. Unlike the conditional or unconditional stimulus involved in classical conditioning, the antecedent event does not elicit responding, but instead signals that the response will have reinforcing consequences (Skinner, 1938). The antecedent event is most commonly referred to as the *discriminative stimulus.* This reflects the observation that an organism under operant control will *discriminate* among stimuli, in that it will respond only to those contexts which have previously occasioned a reinforced behaviour, but will fail to respond to those contexts which have not. An organism can then come under what is known as *stimulus control,* the capacity for an object or event to have a systematic and demonstrable effect on operant behaviour. Conversely, stimulus control may transfer to other stimuli that the organism has not had prior experience with - behaviourists refer to this failure to discriminate as the *generalisation* of learned behaviour. Finally, a contingency may mould behaviour into a particular form through the differential reinforcement of some topographical qualities over others. This process of selectively reinforcing those behaviours, which incrementally approximate an optimal response form, is known as *shaping.* Thus, operant conditioning allows the organism's behavioural repertoire to diversify and expand over the life span.

Central to the school of radical behaviourism is its endorsement of *functionalism.* Functional definitions are those which describe an entity or event in terms of its purpose. In the analysis of behaviour, functionalists asks the question "what role does conditioning serve in the relationship between organism and environment?" For example, classically conditioned responses of adrenergic activity may serve to increase one's readiness for

escape when there is a likelihood of danger. Of course, there are probably times when respondent conditioning is unwanted (e.g., conditioned phobic responses), but classical conditioning per se serves to benefit the species as a whole. Operant conditioning functions to increase an organism's ability to deal with the environment also. For example, a child would not develop in school if she[2] did not repeat and increase those behaviours which earned praise and encouragement (i.e., reinforcement). The advantages of having the capacity to learn are infinite in number, and the ability to gain from past experience is essential in maintaining both survival and quality of life. Similarly, however, the occurrence of learning which is ultimately destructive to the individual is obviously also possible (e.g., learning to take illicit drugs).

I will further elaborate on the importance of functionalism as it relates to an organism's capacity to effectively deal with its environment in the context of the *law of effect* (subsection 1.6). At this point, it is more important for the reader to understand that functionalism as a *philosophy* greatly serves the radical behaviourist at the level of scientific analysis. By definition, an *operant* refers to a class of responses or events that are identical with respect to their functional properties: that is, their effect on the environment (Skinner, 1969). Functional categories are important to a radical behavioural analysis for two primary reasons (Baum, 1994). First, no two behaviours are ever *exactly* alike in a topographical sense, and even if they were, they would have occurred at different times (i.e., they would have differed on at least *one* dimension). Thus, behaviours *need* to be grouped in some way before they can be meaningfully spoken of. Second, there is often nothing to be gained by definitively separating two events in an operant analysis of behaviour. For example, "ringing someone's doorbell" and "knocking at someone's door" may both earn the reinforcement of "being greeted at the door". In the sheer interest of parsimony, there may be no point in distinguishing between behaviours that are functionally equivalent with respect to the contingencies they are bound to, and from a pragmatic point of view, such simplicity may be useful. Functional definitions are in contrast to *structural* definitions, which merely detail the *physical characteristics* of an event, without regard to function. An example of a structural

[2] Feminine pronouns will be used where the sex of the referent subject is ambiguous.

definition might be to describe the behaviour of "knocking at the door" by its physical dimensions, such as speed, total number of knocks, time duration of knocking, and so forth. Later I will provide a more elaborated explanation of the nature of structural definitions, because they too can be very useful within a radical behavioural analysis. However, as we shall see in the following subsection on verbal behaviour, functional definitions are of primary importance to the major themes of this project.

1.2 Verbal Behaviour.

Perhaps the most difficult class of behaviour to study scientifically is the verbal behaviour of humans. One reason for this might be the fact that language is determined by so many variables, seemingly random and devoid of any quantifiable order. Behaviourism, however, does not regard *any* natural events as randomly occurring variables, but simply regards verbal behaviour as complex operant behaviour under the control of environmental stimuli (Skinner, 1957). Of course, the fact that people respond to statements, requests, advice, or any other verbal utterance made by others is not a discovery for which radical behaviourists claim responsibility. Within the realm of behavioural science, however, Skinner was a pioneer in the sense that he believed that verbal behaviours could be understood in the same way as other behaviours: subject to the forces of reinforcement, punishment, stimulus control, and many of the other conceptual tools that have enabled the development of a reputable behavioural science.

Skinner (1957) defined verbal behaviour as "behaviour reinforced through the mediation of others" (p. 14). As it is used here, the term *mediation* refers to the presence of an intervening but not independent variable. Thus, verbal behaviour is a class of operant behaviours, established through contingencies of reinforcement which are in some way regulated by others. Verbal behaviour can be seen as a subset of the wider behavioural category of *communication,* which is simply defined as the generation of stimuli by one organism which alters the behaviour of another (Baum, 1994). According to Skinner, verbal behaviour always involves at *least* two organisms: the *speaker* (generating the stimuli) and a *listener* (who responds). Any one instance of such verbal

interaction can be termed a *verbal event*, and when a speaker and listener repeatedly swap roles in a sequence of verbal events, they are said to have participated in a *verbal episode* (Skinner, 1986). Two or more people who engage in multiple verbal episodes over a period of time constitute the *verbal community*, and it is the members of one's verbal community that provide the reinforcement and motivation for verbal activity (Skinner, 1957).

Verbal behaviour is not necessarily *vocal* behaviour. To paint a sign is verbal behaviour, provided the effect it has on those who read the sign is the mediating consequence of one's "sign-writing" behaviours. Conversely, whistling a tune while alone is probably not verbal behaviour because it does not involve a listener. However, like all operants, verbal behaviours are always only defined by their controlling variables. If a friend had reinforced my whistling in the back yard on previous occasions (e.g., praised it), this would still qualify as verbal behaviour as it was established through the intermittent reinforcement of another (i.e., mediated by my friend). It is also important to note that it may not always be the *original* listener that continues to provide reinforcement for any established verbal behaviour. In fact, once my behaviour of "tune-whistling" had been established, reinforcement might have been mediated by any one of my friends or family who cared to reinforce this behaviour. Thus, verbal behaviour can generalise to different settings where various members of the verbal community may provide further reinforcement.

Skinner (1957) identified many different units of verbal behaviour, several of which I will outline now. A *tact* is verbal behaviour under the control of an environmental object or stimulus property (Skinner, 1957). Reinforcement for tact behaviour is contingent on accurately identifying and responding to a previously learned association between a verbal utterance and a stimulus or its properties. For example, exclaiming "this sandwich is stale" would be a tact if such statements had previously been reinforced in the presence of stale sandwiches. It is as though the presentation of the sandwich itself controlled the likelihood of such an utterance. We learn to tact through earning verbal reinforcement for pairing verbal utterances with stimulus events in a way that is recognised and reinforced by the verbal community. As such, tacts are usually easily tested for their stimulus-verbalisation correspondence. For example, if someone gave me an apple and I said

"thanks for the orange", it is likely that I would quickly receive verbal admonishment from a member of my verbal community. Tacting is probably under relatively tight stimulus control, learned from a young age, and relatively insensitive to extinction (e.g., one might be reinforced for correctly saying "carrot" in the presence of carrots from a young age, and then require little or no further reinforcement to maintain this tact throughout their life). One can tact a very basic or short-lived event, such as a hand-clap, through to the tacting of complex operants of long duration, such as a musical composition (Catania, 1998).

By contrast, a *mand* is verbal behaviour under the control of the reinforcing effects of altered listener behaviour. Thus, mands are verbal behaviours controlled by the state of reinforceability in the speaker (i.e., relative states of deprivation or aversive stimulation alleviated by listener responses) and the extent to which the presenting stimulus indicates an availability of reinforcement from the listener (Skinner, 1957). To continue with the first example, saying "is this sandwich fresh?" in order to receive agreement from others, largely irrespective of whether it was a fresh sandwich or not, would be a mand. It is as if the availability of an answer from the listener controlled the likelihood of such an utterance. As a response class, manding generalises to a wide range of situations in which such behaviour is likely to alter the behaviour of another in a way the listener finds reinforcing. The person in the example may have never before sought an answer about "fresh sandwiches" from those particular people, but because of a long history of being reinforced with agreement for such statements, this behaviour could generalise to other people who might also recognise and respond to "statements about the freshness of food".

It is important to acknowledge that the tact and the mand are defined by their functional properties, and not by their structural characteristics (e.g., their exact words or linguistic composition). It is entirely possible that the statement "this sandwich is stale" could be mand, if the controlling variable for such behaviour were the likelihood of a desired listener response. It is also possible that the statement "this sandwich is stale" is a tact *and* a mand, provided that the functional criteria for both had been met. The distinction is not complicated if one simply understands that the reinforcement of achieving a desired effect on the listener is separate from a contingency of reinforcement for correctly producing a verbal utterance which has been experientially paired with a

stimulus[3]. Thus, the tact and the mand may sometimes *appear* the same, but their controlling variables are different (Catania, 1998). Nevertheless, tacts probably most often appear as descriptions, statements, and assertions, while mands commonly resemble requests, appeals, and inquiries.

Skinner (1986) identified two major types of mands. An *object*-mand is a mand controlled by an apparent availability of the item specified as reinforcement. For example, the statement "pass me the butter" might be reinforced with receiving a block of butter from the listener. Alternatively, an *action*-mand is a mand controlled by an apparent availability of the listener's *activity* as reinforcement. For example, the statement "pass me the butter" would be an action-mand if the reinforcing agent was seeing the listener pass the butter, rather than the actual receiving of the butter itself. The reader may note that there are elements of tacting in both of these mands, in that the speaker has tacted either a behavioural unit or an object, and components of these two mand statements could qualify as tacts (e.g., the word *butter* appears to have been correctly tacted). Once again, this illustrates the fact that many instances of verbal behaviour and their components can arise from *multiple causes* (Catania, 1998). Thus, the same formal unit of verbal behaviour can often be conceptually broken down into multiple functional units.

Two other units of speaker behaviour, distinguishable from the tact and the mand, are the *autoclitic* and the *intraverbal* (Skinner, 1957). Verbal behaviour that depends on other verbal behaviour of the speaker for its occurrence, and modifies the function of that behaviour, can be termed autoclitic (Catania, 1998). Take the statement, "I want *you* to take me home right away". In this case, an autoclitic might be the accentuation of the word *you*, as it changes the function of the statement "I want you to take me home right away" so it can operate as though the speaker had said "I want you to take me home as opposed to any one else doing so". A similar accentuation could alter the function of this statement in the same way, such as "I *want* you to take me home, or "I want you to take me *home*". Thus, the tonal accentuation of certain words is a relatively flexible class of

[3]As I will detail later, Skinner devised the term *impure tacts* for such units. An impure tact is a tact which is *partially* controlled by a state of deprivation or aversive stimulation in the speaker, or partially controlled by an apparent availability of reinforcement by the listener (Skinner, 1957).

autoclitic behaviour. However, autoclitics take a diverse range of forms. For example, saying "I want you to take me home!" at a greater volume might also serve an autoclitic function. An autoclitic can also involve the addition of a whole word, such as the immersion of the word *don't* to form the statement "I *don't* want you to take me home". By definition, an autoclitic can not stand alone because it needs to be attached to other verbal behaviour for its effect. For example, the word *can* is in most cases used as an autoclitic. It is very unlikely that the word *can* would ever be used in isolation, in the form of a sentence. However, preceding the sentence "dogs eat grass", the word *can* could transform the function of this statement into a question (i.e., "*can* dogs eat grass?"). Thus, an autoclitic is verbal behaviour which is dependent upon the preceding or ensuing occurrence of other verbal-behaviour for its desired effect on the listener (Skinner, 1957).

However, some verbal behaviour may depend on the presence of other verbal behaviour of the speaker for its occurrence, but not necessarily alter its function. An *intraverbal* is simply verbal behaviour which is in response to (or cued by) other verbal behaviour of the speaker (Skinner, 1957). Intraverbals are perhaps most easily illustrated by considering long chains of intraverbal behaviour. Reciting the alphabet, for example, is a sequence of intraverbal responses. Saying "C" after "B" is not a tact of "C", because saying "C" in this manner has not come about through the general verbal community reinforcing "C" only in the presence of "B". Rather, "C" is a response to "B" only when engaging in the intraverbal chain of alphabet recital. An intraverbal is only linked to stimuli through thematic verbal chains, and unlike the tact, its correspondence with the stimulus is only arbitrary rather than experiential (Catania, 1998). We may engage in long sequences of intraverbal behaviour, such as singing a song to oneself or privately calculating some problem. Alternatively, we may engage in short intraverbal sequences – the word *go* in the sentence "ready, steady, *go*!" is intraverbally derived. The tact, mand, autoclitic, and intraverbal are all very important functional units of speaker behaviour.

1.3 Problem-Solving and the Concept of RGB.

According to Skinner (1984), a problem situation is any contingency for which there is no immediately available response. These are most likely to be complex contingencies

which demand responses in a refined or composite form. A fundamental premise of Skinner's position on problem-solving is the notion that the *speaker and listener can be the same person* (Skinner, 1969). As such, people may speak or behave in ways that will allow them to respond to their own behaviour. Thus, a behavioural event of the speaker can also act as a stimulus to the speaker[4], and verbal behaviour then becomes *behaviour mediated by oneself.*

Skinner wrote extensively on the role of self-mediated verbal behaviour in human problem-solving (see Skinner, 1984). *Precurrent behaviour* is behaviour which generates a discriminative stimulus which then evokes a reinforceable response. Sequences of precurrent behaviour may enable one to find an effective solution to a problem. For example, if I were faced with the problem of a broken car engine, the response required to fix it may not be readily apparent, or even possible. In fact, faced with this dilemma, I might have to remove the head gasket, the engine valves, and so on; a chain of precurrent behaviours which would each generate the stimuli needed to evoke the next response required to help solve the problem. This would presumably continue until the broken piece or displacement was found, and a solution would then be available. Thus, chains of precurrent behaviours (and responses to the stimuli they produce) allow the process of problem-solving to occur. However, Skinner (1984) would argue that there is another important aspect to this example in a verbal sense; the stimuli generated by my precurrent behaviours could be used by some *other* member of the verbal community to solve the problem. I might cease trying to fix this engine when the problem was not fully solved, but somebody else could then continue working on that same engine. Thus, the stimuli generated from precurrent behaviour can exert control over the behaviour of other members of the verbal community who share a common interest in the same problem.

There is another way a speaker can assist another member of the verbal community to respond to a contingency. To continue the previous example, if I wanted to show someone else how to fix a similar engine problem, I might draw a diagram showing the sequence of behaviours needed to successfully conduct the reparation. My diagram may also include specification of the original symptoms of the car, such as stuttering or excess

[4] In such cases, the speaker has become a *listener* of her own behaviour.

emission of exhaust fumes, and perhaps the particular engine model. This would enable anyone reading the instructions to recognise whether their own car had the same fault. With the use of the diagram someone else could then fix any number of cars with the same fault, without having to go through the time-consuming process of trial-and-error that was required when I had fixed the engine. Skinner (1969) would argue that the behaviour of someone following such a diagram could be termed *rule-governed*, evoked or occasioned by the statement of a *rule*. Skinner (1969) defined a rule as a "contingency-specifying stimulus". This means that any rule must imply three things: (a) a setting or stimuli for the behaviour, (b) the actual behaviour required, and (c) the expected outcome or reinforcer. Someone following this diagram of mechanical repair would be responding to a rule because it specified a context or set of stimulus conditions (the original symptoms and particular type of engine), the behaviour (the exact sequence of behaviours required to fix it), and the potential reinforcement for these acts (a running vehicle).

Rule-governed behaviour (RGB), like all verbal behaviour, requires both a speaker and a listener - someone to state the rule and someone to follow it (Baum, 1994). If a speaker says "go and get the mail", the three components of a contingency are being specified to the listener: (a) the context (going outside), (b) the behaviour (looking in the mailbox), and (c) the reinforcement (e.g., receiving the mail). A rule can not function if it does not in some way imply each of these three elements. Here to say *imply* is to say "allows the recognition of". For example, in the case above, the listener may not have been familiar with the stimuli specified in the rule, such as the location of the letterbox. In this case, the speaker may need to add components to the rule, perhaps physically pointing towards the place of the letterbox. Even the most basic verbal units can act as rules. For example, the rule "stop that!" could imply the three aspects of a contingency if the listener's (or *rule-followers*) experience with the rule was sufficient enough to allow compliance. A parent who issues the command "stop that!" may be administering a rule that functions exactly the same as "stop what you are doing right now or I will punish you". However, the three aspects of a contingency needn't be *structurally* embedded in the statement in this way for it to qualify as a rule. Rules are simply a functional class of stimuli, just as food or books are. What is unique about the function of this class of

stimuli is that they *direct* listeners in responding to a contingency, rather than from a prior history of that contingency shaping the listeners behaviour in and of itself.

As speakers, humans tend to reinforce RGB so that it occurs again, because having the ability to specify a contingency to someone else has obvious advantages. Imagine if the speaker in the last example was unable to get the mail (e.g., perhaps the speaker was bedridden). Waiting for the natural contingencies surrounding the listener's "mail-getting" behaviours may take years to fully induce the behaviour of fetching the mail - it may never even occur without some type of instruction. As listeners, humans respond to rules because "rule-following" has usually earned reinforcement in the past, which has often been provided by the original speaker of the rule. A common but invalid objection to this explanation is the observation that non-verbal organisms often respond to what appear to be contingency-specifying stimuli. For example, someone may say to their dog "get the paper", and the dog may do so. However, the dog is only responding to the formal qualities of the event, rather than any recognition that this utterance specifies any new contingency. That is, the dog's behaviour of getting the paper would need to be repeatedly reinforced following the presentation of this statement before it could begin to act as a discriminative stimulus for "paper-getting" behaviour. It is unlikely that any dog would be able to get the paper for the first time on hearing such a rule, yet most humans would be able to, recognising that the rule *specified* an instruction for which compliance with would result in reinforcement.

Rule-following is a highly generalised class of responses in normally functioning adults. According to Reese (1989), this generalisation takes at least three main forms: (a) the following of rules for which compliance has earned reinforcement, generalising to the following of new rules which share similar properties, but have not yet produced a reinforceable response, (b) behaviours occasioned by self-produced rules for which compliance has earned reinforcement, generalising to the following of self-presented rules which have not yet produced a reinforceable response, and (c) following rules provided by people who have not stated rules to the listener previously, generalising to new rule-givers who share similarities with people who have previously presented productive rules (e.g., people with notable authority). Thus, we not only discriminate between rules and classes of rules according to the consequences they indicate (e.g.,

identifying dangerous rules), but also between the contextual stimuli that accompany them, such as the people and types of people that give us rules (e.g., distinguishing between rules presented by polite salesmen versus those presented by rude salesmen).

Verbal episodes in which the speaker and the listener are the same person can occur at the covert level. As mentioned earlier, Skinner (1974) emphasised that covert verbal behaviours are similar to public events that can function as discriminative stimuli. In fact, Skinner often made the point that some forms of thinking can be described as "talking to oneself". Of course, private verbal behaviour does not always have to involve *words*. One might whistle a tune "in their head", and this might be reinforced by the satisfying feeling produced by listening to music. Thoughts, self-talk, and any other events only observable to the individual actually experiencing the activity are as scientifically valid as events which are directly accessible to the public. As such, we can understand how humans could present rules to *themselves*. One might say "I need to go and get the mail now" and subsequently do so. Thus, as a listener, this person would have followed her own rule. People quickly discriminate between the form and shape of covertly recited rules which evoke unreinforced behaviours from those that evoke reinforced behaviours, and private rule-formulation behaviours quickly evolve when one successfully generates rules that provide even small amounts of reinforcement.

People formulate rules by recognising contingencies - situations which act as discriminative stimuli for the precurrent behaviour of rule-recital. A rule may not accurately describe a contingency, but for it to function as a rule, it does not have to. We may create rules that do not even correlate with some real-world contingency, but they may still function as rules. For example, take the rule "If I jump off the Eiffel tower I can fly to Mexico". This would not be a rule worth following, but it would still have been created in response to some set of conditions that stimulated the speaker to formulate such a rule (e.g., perhaps this speaker was delusional, and derived the rule from looking at a picture of a bird flying off the Eiffel tower and incorrectly applied it to herself). Creating and presenting oneself with a rule is a type of precurrent behaviour because it involves the generation of a verbal discriminative stimulus which can serve to evoke reinforceable behaviour. The more effective a rule proves to be, the more likely we will generate subsequent rules that contain *components* of that rule - perhaps by identifying

the effective elements of rule-formulation. Alternatively, we may simply try to solve different problems with the same reliable and trustworthy rules. For example, upon first meeting someone I wish to establish rapport with, I might follow my well-used rule "shake someone's hand and smile when you first meet them and you will make a favourable impression". I might use this rule because it had worked many times before, and in many novel social situations. Because we can recite rules given to us by others, the original provider of the rule does not have to be present when rule-following occurs, even if we have not had first-hand experience with the contingency specified in the rule. One may even recite a rule out loud when engaged in a task so as to benefit from the extra stimulus control that audible rules can exert. Of course, we usually do not do this because talking out loud to oneself can result in socially-administered punishment. However, it is interesting to note that children, typically less aware of such social sanctions, often talk to themselves when completing a complex task, perhaps for this very purpose (Roberts & Tharp, 1980).

The verbal community constantly engages in the trading of effective and useful rules. Perhaps the labour of a teacher might be seen as that of providing effective rules for children, in addition to shaping their non-verbal behaviour. We may present our novel rules to someone else, as this can earn reinforcement too - people tend to reinforce the behaviour of those who present them with valuable rules because it increases the chances that the speaker will expose them to *more* useful rules on later occasions. Conversely, as listeners we tend to punish the verbalisation of poor or misguided rules so as to decrease the chances that the speaker will repeat them to either themselves or others whom we care about.

Skinner frequently made a distinction between RGB and contingency-shaped behaviour in order to illustrate the importance of rule governance. Strictly, rule following is also contingency-shaped behaviour, but what stands RGB apart is that the primary controlling stimulus is always the rule itself (Baum, 1994). There may of course be an array of reinforcers that act upon the behaviour evoked by any particular rule, but the stimulus control of that rule has *been acquired through repeated instances of reinforcement for rule-following*. It is this type of reinforcement that maintains RGB. Any one instance of rule-following may or may not be reinforced by a member of the

verbal community, but as a deeply ingrained class of verbal behaviours, rule-following does not depend on a continuous schedule of reinforcement to be maintained. As opposed to RGB, contingency-shaped behaviour is never stated or spoken about to the listener in a way that allows recognition of the contingency a priori. Take Skinner's (1969) example:

> An outfielder moves to catch a ball. Following its trajectory, he moves under it and grasps it with his glove. This event is undoubtedly contingency-shaped. The outfielder is simply responding, as he has done hundreds of times, to the effects his behaviour has on moving toward the ball. A dog can acquire the same behaviour in the same way. Contrast this with the ship captain moving to 'catch' a descending satellite. The trajectory of the satellite is analysed in detail. Mathematical models are consulted which take into account a host of factors such as wind-speed and drag coefficients. Its place of impact is predicted and approached. This behaviour is not controlled by the past consequences of trying to catch satellites – it is controlled by *rules* (p. 146).

Although apparent in Skinner's (1969) example, the distinction between RGB and contingency shaped behaviour is not always as observable. Behaviour may first occur as rule-governed, and then *become* contingency-shaped once in its final form. For example, take the rule "push the accelerator slowly and you will accelerate smoothly". Presenting such a rule to first time drivers would rarely results in a flawless takeoff. Although initial attempts would probably *approximate* this rule, the final form is most probably controlled by unspoken contingencies provided by the "feel" of the car, such as the gearbox grinding and jerky movements. This may be one of the most useful functions of RGB because it enables speakers to provide an initial cue for a listener to respond to a contingency, which can then be maintained by the natural contingency-governed consequences of responding. Conversely, behaviour may begin as contingency-shaped but then come under the control of rules. For example, a coach may give directions to an outfielder so as to improve the efficiency of her catching style. Individually, someone may already be responding to a contingency, but formulate a rule so that "he himself can

then react more effectively either now or later when the contingency-shaped behaviour has weakened" (Skinner, 1969, p. 159).

I have argued that rules exert stimulus control in the same way that non-verbal stimuli do, although this has been the centre of much debate in recent years. RGB often occurs *straight after* the rule has been presented by another speaker or recited by oneself. In these cases, a rule evokes behaviour in a fashion consistent with traditional notions of operant conditioning, in that a stimulus has *occasioned* a response. However, the way in which rules are seen to exert stimulus control is sometimes seen as contradicting this definition. A person may be told a rule *years* before that statement eventually causes an alteration in behaviour. For example, a parent may tell her child, "if you ever see a man pull out a gun, try and hide", and the child may not encounter a situation where she could follow this rule until many years later. This is not in accordance with our usual reference to rules as discriminative stimuli, which are said to evoke an *immediate* but momentary change in behaviour due to a special history of differential reinforcement. Even from the outset, Skinner (1969) defined operant conditioning as "the increased probability of a response occurring in the *presence* of stimuli for which that response was on prior occasions reinforced" (p. 133).

Thus, in cases where rules have not immediately preceded the behaviour they are said to have caused, there appears to be no clear means of explaining this control in terms of normal discrimination theory (cf. Cerutti, 1989). Baum (1994) has proposed one way in which we might view this conceptual problem, which he refers to as the *gap-in-time dilemma* (reflecting that the logical dilemma concerns the irregularity of having an extending delay between the stimulus rule and the rule-governed response). Baum (1992) argues that these gaps in time present no dilemma for a radical behavioural analysis because the rule still *technically* exerts stimulus control when temporally distant from the behaviour of concern, in that this control can still be measured as an *observable relationship* between environment and behaviour. Baum's solution is essentially an appeal to the *correlation-based law of effect*. The correlation-based law of effect holds that two sets of natural phenomenon can be seen as having a causal relationship if they consistently occur as a function of one another, regardless of any spatial or temporal contiguity (Gale, 1979).

However, the solution offered by Baum (1994) does not seem to *explain* how a physical event can affect another temporally or spatially removed event, but merely argues that such an explanation is *unnecessary*, and simply avoids providing a comprehensive description of all the determinants in the sequence of events that might allow this type of "action-at-a-distance". Such a description can be demanded of radical behaviourism because it endorses *determinism*, the contention that all events are the direct result of those immediately preceding it. Appealing to the correlation-based law of effect simply draws attention away from those determinants that might allow temporally removed rules to control behaviour, the most important of which may be private events. If such "hidden determinants" could increase our understanding of behaviour, one can understand why Skinner (1945) urged behaviour analysts to concern themselves with private events for the sake of a *complete* analysis.

This thesis does not subscribe to the notion of non-contiguous stimulus-response causation, and where a rule is presented and evokes compliance after a significantly delayed period, it is assumed that a self-rule, most probably recited covertly, has transferred to a new situation as generalised precurrent behaviour (i.e., so it *can* then act as an immediate verbal antecedent). This is therefore in accordance with Skinner's (1969) definition of a discriminative stimulus as physically occasioning the behaviour it has evoked. Such adherence to Skinnerian theory only seems apt given that this is also where the notion of RGB originated. More importantly, however, because the notion of RGB is dependent on laboratory research (as this is where the idea of discriminative control is validated) it necessary to be diligent in confining ourselves to a usage of the term which is consistent with that research. And in the non-human laboratory, discriminative stimuli are never temporally or spatially removed from the behaviours which they occasion (Schlinger, 1993).

1.4 Functional Units of RGB: Tracking and Pliance.

We are now ready to divide RGB into several functional units. Baum (1994) has identified two distinct sources of reinforcement for RGB, that provided by the *proximate*

contingencies versus reinforcement provided by the *ultimate* contingencies. Baum distinguishes RGB from simple contingency-shaped behaviour by pointing to the fact that RGB always involves the consequences of both an immediate source of social reinforcement for rule-following (proximate reinforcement) and a long-term or distant source of reinforcement for responding to the rule (ultimate reinforcement). For example, a parent may say to her child "put on your shoes" and the child may do so to avoid the aversive consequences mediated by the parent for non-compliance. This would be contact with the proximate contingency, the most immediate and obvious incentive for this instance of RGB. However, presumably the child's behaviour of wearing shoes would also be in contact with a contingency which provided intrinsic benefit to the survival of the child, such as the avoidance of foot injury or disease. This, according to Baum, would be the ultimate contingency, and is essentially the core reason that humans use rules in the first place.

Baum (1994) makes a valid distinction between these two types of reinforcement, but the current thesis argues that he has made an error in assuming that *all* cases of RGB must involve both a proximate contingency *and* an ultimate contingency. For example, there are probably instances of RGB which are only in contact with the proximate contingency. For example, take the rule "educate yourself, and you will be rewarded". School-goers may follow this rule to avoid the consequences mediated by their parents for non-compliance. But there may be no ultimate contingency involved - the rewards for educating oneself may never actually occur, regardless of the fact that most parents would administer such a rule so that their children could *ultimately* benefit from the advantages of being educated.

Nevertheless, Baum (1994) rightly points out there are two major reasons why someone would follow a rule. First, someone might follow a rule because it is clear that the person giving the rule is going to arrange consequences for compliance or non-compliance, what we might call an arbitrary or *contrived* contingency. Second, someone might follow a rule because it specifies the availability of reinforcement from a contingency that the speaker has no control over (i.e., is independent of) - what we might call an environmental or *natural* contingency. For example, someone might say to a listener, "do not open the can with a knife, use a can opener". There would be two

obvious contingencies which could evoke the following of this rule. The natural contingency (what Baum might refer to as the ultimate contingency) might be the negative reinforcement of "not cutting oneself", as people are prone to do when opening a can with a knife as opposed to using a can opener. This contingency is mediated purely by the environment, and not verbal in any strict sense. Second, this listener might simply follow this rule because it was clear that the speaker was going to arrange consequences for doing so. Perhaps the listener was a young child - unable to understand the drastic consequences of being cut, the child might simply use a can opener instead of a knife to avoid being reprimanded by her parents. In fact, the parent who administered the rule might provide reinforcement by saying "Good, see how you didn't cut yourself?" Zettle and Hayes (1982) have also employed this distinction:

> What is crucial here is that RGB always involves two distinct sets of contingencies. One set of contingencies involves those directly related to the behaviour of interest. The second set is verbal and somewhat independent of the first (p. 78).

However, unlike Baum (1994), Zettle and Hayes (1982) use the distinction to divide RGB into its two main functional types, what they have labeled the rule-governance of *tracking* and *pliance*. Natural contingencies most often influence RGB when the rule formally details the non-verbal contingency. For example, the rule "use a knife or you'll cut yourself badly" is likely to induce RGB which is primarily controlled by the consequences of cutting oneself. Thus, in these cases it is the apparent accuracy of the rule to the listener that determines whether it will produce RGB, or stated somewhat differently, the *apparent correspondence between the rule and the natural environment.* Zettle and Hayes named this type of RGB *tracking.* Defined, tracking is RGB primarily dependent on an apparent *environmental accuracy* of the given rule, and such rules are called *tracks.* For example, if a listener followed a diagram of automotive repair because of an apparent correspondence between the diagram and an engine with a problem more or less identical to that specified in the diagram, then this would be tracking. As a response probability, the likelihood of tracking is independent of the likelihood or nature of any speaker-mediated reinforcement for rule-following. Tracking can be considered to

be analogous to the speaker unit of tacting. Like the experiential training of tacts, tracking occurs through a history of being reinforced for correctly responding to a learned correspondence between two events. With tracking however, the correspondence of interest is between the rule and the natural contingency that the speaker is faced with. In this example, the listener might have learned to track automotive rules by effectively responding to a correspondence between "engine repair diagrams" and "broken engines" by emitting appropriate "engine repair behaviour". If, for example, there was a notable conflict between the rules of repair and the broken engine, then tracking might be less likely to occur (e.g., perhaps the symptoms of the engine in the diagram did not match those of the present engine, indicating that the instructions were unsuitable).

Alternatively, contingencies for compliance which are mediated by the *speaker* of the rule can also determine the likelihood of RGB. Such rules most often formally specify the speaker-mediated consequences available to the listener for compliance. For example, if a child was given the rule "if you use a knife instead of a can opener you will be going to bed early", it would probably be followed in order to avoid the speaker-administered punishment of being sent to bed. Zettle and Hayes (1982) have termed this type of RGB *pliance*. Defined, pliance is RGB controlled by an apparent[5] availability of speaker-administered reinforcement contingent on following the rule. Rules which induce pliance are termed *plies*. As a response probability, the likelihood of pliance occurring is controlled by an apparent correspondence between the rule and the speaker-mediated consequences it predicts. For example, if a listener were to follow a diagram of automotive repair because it was apparent that the speaker who constructed the diagram was prepared to administer consequences to control this behaviour, then this would be a form of pliance, and the diagram would have served as a ply. Note that with pliance a listener could follow such a diagram yet be relatively indifferent to the accuracy of the rule, that is, whether the diagram *actually would* fix the car. Pliance can be considered to be analogous to the speaker unit of manding. Like a mand, pliance occurs through a history of verbally-mediated reinforcement for correctly responding to a correspondence

[5] Here I use the term "apparent" to denote the fact that the likelihood of responding is dependent on the listener's history of reinforcement with similar rules, independent of the social or naturally programmed contingency.

between the verbalisation and its potential effect on the listener. As such, Zettle and Hayes argue that pliance is likely to be dependent on such variables as the authority of the listener in relation to the speaker (e.g., a mechanic giving such a diagram to an apprentice would be likely to produce pliance), the availability of the resources that the speaker needs to reinforce rule-following (e.g., perhaps it was obvious to the listener that the speaker could not pay the listener to follow the diagram, in which case pliance may not occur), or anything else that had some kind of bearing on the apparent likelihood of speaker-mediated consequences.

It is important to note that not every *instance* of pliance will require the administration of speaker-administered reinforcement for it to be defined as so, because pliance, like many other response classes, can be maintained purely through intermittent reinforcement. In order to meet the definition of pliance, the evoked RGB must simply be a function of contingencies surrounding rule-following per se, either provided by the speaker or similar speakers on prior occasions. This is a very important point because throughout this thesis I will refer to single instances of pliance where there is no obvious likelihood of the speaker administering reinforcement. However, as a general class of stimuli, plies only need to have been previously associated with social reinforcement for rule-following in order to evoke pliance.

However, to use the term *social* in the context of pliance is to only refer to sources of reinforcement which have been conditioned with the presence of a speaker or other rule-givers, not just the considerably broader category of "human-mediated reinforcement". For example, take the rule "Men act favourably to women who reassure them of their masculinity". If a woman were to follow this rule in order to reap the consequences that a man might mediate for such praise, then this would *technically* have been socially-mediated reinforcement (men, after all, are humans or *social* beings). However, to follow this rule would not be pliance, because the speaker of the rule was in no way causally connected to the administration of such reinforcement, neither as a personal discrimination or as a generalised stimulus. Thus, reinforcement mediated by "men" for following this rule would be an environmental contingency, and to follow this rule because it appeared to accurately correspond with that contingency would be an example of *tracking*. The key point to remember is that when a person follows a rule because

another person has differentially reinforced varying approximations of compliance in *connection* with that rule, then pliance has occurred.

Zettle and Hayes (1982) have also extended their concepts of tracking and pliance to incorporate *self*-produced and followed rules. As mentioned before, an individual may formulate rules about contingencies that they have had contact with in order to prepare for similar contingencies, or may simply construct rules because of the reinforcement provided by the individual's verbal community for doing so (e.g., the reinforcement someone might receive for giving advice on how to effectively respond to a contingency). Zettle and Hayes agree that for someone to follow their own self-presented rules, the speaker and listener must be the same person. If *self*-mediated reinforcement controls the following of a self-administered rule, then this would be a type of *self-pliance* (Zettle & Hayes). For example, someone might recite the self-rule, "If I tell my boss that I disagree then I can thank myself for being assertive". If this person had planned to make "thanking oneself" contingent on expressing a difference of opinion to her boss, and the salience of this was a determining factor for the following of the rule, then this would be a self-ply. That is, "thanking oneself" would have been speaker-mediated.

It should be noted that the notion of self-reinforcement is often considered a dubious one, and some have argued that self-reinforcement is different to normal operant conditioning in several crucial ways (see Goldiamond, 1976). For example, the consequences involved in self reinforcement are often not truly *contingent* on the emission of a response, as evidenced by occasions where people administer themselves with reinforcement without having executed the required behaviours; what Martin and Pear (1996) refer to as *behavioural short-circuiting*. Thus, although self-reinforcement is probably effective in many circumstances, throughout this essay there will be a consistent effort to consider other causal variables that do not involve an appeal to self-mediation.

Accordingly, self-pliance might be better understood in terms of its connection with reinforcement previously administered by an *external* listener (i.e., a speaker other than the listener). In fact, as listeners we can instruct speakers to give themselves rules, and then reinforce a correspondence between the new self-rule and the behaviour it evokes; that is, reinforce the accuracy or extent of compliance (Zettle & Hayes, 1982). For example, the person in the present example may have been told many times before to

"stand up for themselves", and then received verbal praise from the speaker upon compliance (e.g., perhaps a speaker of this rule had once said "great, I knew you could be assertive if you tried!"). On a later occasion, the speaker may then recite the same rule when confronted by her boss, and subsequently follow it. Self-mediated or external reinforcement may not be administered, but the rule would still have been a self-ply if it was prior experience with contingencies for rule-following per se from which the rule's stimulus control originated.

Alternatively, if an apparent availability of reinforcement from the natural contingency implied in a self-rule determined the likelihood of a response to that rule, then this would constitute a *self-track*. (Zettle & Hayes, 1982). For example, someone might say to themselves, "It's about time I simply told my boss that I disagree", but follow it in order to avoid the long-term natural consequences of being passive and unresisting, such as diminished respect from the person's employer. If the availability of this reinforcement was apparent to the speaker, and this evoked the speaker's RGB, it would be an example of *self-tracking*.

Before continuing with our discussion of RGB, the reader should be aware of several issues. First, it is true that we have now made substantial additions to Skinner's original definition of a rule as simply being a "contingency-specifying stimulus" (Skinner, 1969). This definition of Skinner's did not account for the distinction between the social versus the natural contingencies that maintain rule-following, although Skinner (e.g., 1974) often gave examples where the two could be seen to be controlling RGB in different ways. It should now be apparent that here the distinction between the natural versus the social determinants of RGB is used as a defining characteristic of any case of RGB. Second, although I have employed Zettle and Hayes' (1982) concepts of tracking and pliance, there is one aspect of their definition which is not subscribed to here: Zettle and Hayes, like Baum (1994), state that by definition RGB must as always be in contact with both a social *and* a naturally-mediated contingency. In fact, Zettle and Hayes (1982) define RGB as "behaviour under the control of two contingencies, one of which includes a verbal antecedent" (p. 78).

In the underlying quote, Zettle and Hayes show how simple contingency-shaped behaviour can be distinguished from RGB by giving an example in which RGB is

necessarily under the control of both a rule (which has acquired stimulus control through verbal interaction) *and* a red light (which has acquired discriminative properties through direct contact with similar non-verbal contingencies):

> Consider a complex operant task which is presented to two different subjects. One subject is merely presented with the task with no further instruction while the second subject is told by the experimenter "If you perform responses a, b, and c, following lights x, y, and z, you will be rewarded with money". In time, the behaviour of both subjects in the task will appear quite similar if not identical, although the controlling variables would be quite different. The first subject's behaviour (setting aside, for the moment, self-generated rules) is under the control of the lights. The behaviour of the second subject is under the control of the rule *and* the lights...It can be seen that all effective rules are discriminative stimuli, but not all discriminative stimuli are rules (p. 77).

However, this is an imprecise definition, and the same example given by Zettle and Hayes (1982) can be used to illustrate its deficiencies. Without knowing all of the relevant factors (e.g., the private verbal behaviour of the subject), it is impossible to exclude the possibility that the responses of the second subject were sometimes *entirely* controlled by the given rule "do x, y, and z". At other times, the second subject may have emitted the responses of "a, b, and c" *purely* under the control of the red lights. Thus, *defining* RGB as only those responses under the control of both a rule *and* a non-verbal stimulus is excessively narrow, and may severely restrict the utility of the concept. Of course, there may have been instances in the current example in which the subject's response actually *was* under the control of both the rule and the lights, but this is simply another example of *multiple causation*, as referred to earlier, and does not warrant the invention of any new terms or concepts.

Incidentally, it is probably an appropriate time to briefly return to the issue of multiple causation as it relates to the concept of RGB. Just as the same verbal unit can functionally be both a tact and a mand, so can some rules be both a track and a ply. For example, a father may say to his son "take your car to the mechanic". The son may do so

to avoid the negative consequences of disobedience which are mediated by the father; the type of reinforcement that characterises pliance. However, the son may also take the car to the mechanic in order to avoid the natural consequences of not maintaining regular service checks, a rule followed because of its apparent correspondence with the way the non-verbal world is arranged and a classic example of tracking. Thus, we can see how the same instance of RGB can be composed of more than one functional unit, and therefore involve both tracking *and* pliance. But this in no way affects the validity of the conceptual distinction between these two types of RGB. They are different functional units simply because they are controlled by different variables. However, it is true that in Zettle and Hayes' (1982) analysis of RGB, there is usually only reference to examples in which either an apparent speaker-mediated contingency *or* an apparent non speaker-mediated contingency has clearly been the primary factor responsible for RGB. This is done simply because, as we shall see later, most RGB in day-to-day life *is* primarily dictated by either social *or* natural consequences (i.e., RGB is usually quite clearly either tracking or pliance). The point is that this is not definitional, but only the usual state of affairs. The implications of multiple causation will become clearer later in the text, when I discuss how formal versus functional misinterpretation can affect human interaction. For now, however, the reader need only be aware that *pliance* and *tracking* are core units of RGB.

1.5 Empirically Studying RGB

Perhaps the strongest case for the theory of RGB has been made in the laboratory. The advent of cognitive psychology during the early 1970's provoked a surge of experimental activity, and this paralleled a growing interest in the use of instructional variables in behavioural analysis. As Vaughan (1989) has noted, most of the research emerged shortly after Skinner began to focus on works dedicated entirely to the concept of RGB, such as in various sections of his book, "Contingencies of Reinforcement" (Skinner, 1969). This was in contrast to previous works of his that had only contained the concept of RGB as imbedded in the discussion of other topics, most commonly being

essays on problem-solving and verbal behaviour. To date, research on RGB has greatly extended Skinner's initial speculations.

According to Plaud and Newberry (1996), there are three general methods used to study RGB. The *developmental* approach attempts to trace the relationship between the acquisition of verbal control and rule-governance by exposing children of different ages to identical contingencies, sometimes using animals as comparison subjects. The *self-report* method involves asking subjects to describe the contingencies that they have responded to, and these reports are then used to infer generalities about rule-formulation and manipulation. *Rule-provision* involves simultaneously presenting a rule with a contingency and then evaluating the effects of rule-governance by making alterations to the programmed contingency. Most experiments incorporate a combination of these approaches. A term used frequently in these studies is that of schedule *sensitivity* or schedule *insensitivity*. Responding is said to be insensitive when it fails to adapt to alterations in the contingency with which it is in contact (Newman, Buffington, & Hemmes, 1995). For present purposes, however, I will refine this definition by describing insensitivity as *a failure to maximise the amount of reinforcement earned and minimise the amount of punishment incurred, while minimising the extent of response behaviour not directly contributing to an earning of reinforcement.* I will first describe several experiments on RGB and human development which both demonstrate the cogency of speaking about RGB as operant behaviour and provide evidence of several replicable features of instructional control[6]. I will then move on to the phenomena of schedule sensitivity and related features. Finally, research assessing the difference between performance-specific and general rules will be described. I will then summarise the results and implications of this literature.

Developmental research has provided much in the way of showing that RGB is an important and uniquely human phenomenon. Central here is the idea that studying the acquisition of verbal control can be used to monitor the developmental changes that relate to RGB. Interestingly, much of the earliest operant research with humans in the 1950's attempted to *minimise* the confounding effects of verbal control in order to increase

[6] In contrast to several authors (e.g., Catania, 1998), here the term *instructional control* is used synonymously with reference to behaviour governed by *rules*.

experimental consistency and to obtain more "animal like" responding (Vaughan, 1989). However, it was soon accepted that humans simply respond differently to animals in laboratory conditions. For example, humans often fail to show reliable scallops on fixed interval (FI) schedules, post-reinforcement pauses (PRP's) on fixed ratio (FR) schedules, and efficient or appropriate changes in responding when schedule changes are made (Galizio, 1987). Thus, humans are unique in that they do not show the characteristic orderly and replicable patterns of responding which are common within and across non-human species.

Preverbal humans, however, tend to perform similarly to animals. Lowe, Beasty, and Bentall (1983) studied the performance of two infants on several FI schedules. The infants, aged 9 and 10 months old, were each presented with a metal cylinder, the physical tapping of which could earn reinforcers such as music, bread, and fruit. Initially every response was reinforced, but over many trials the programmed contingency gradually progressed through several FI schedules, ranging from 10 to 50 seconds. The results showed a predominant response pattern of a pause in responding after reinforcement, followed by a gradual acceleration in response frequency. The infants failed to show an *insensitivity* in their responding. As mentioned before, *schedule insensitivity* refers to a subject's perseverance with a behavioural pattern in spite of a change in the schedule conditions. As it relates to FI schedules, sensitivity can in part be calculated by measuring the extent to which an interval increase produces three things: (1) an increase in the mean duration of PRP's, (2) a decrease in *relative* PRP's (i.e., PRP expressed as proportion of schedule value), and (3) a decline in *running rate* (i.e., response rate calculated without taking the PRP into account). Both of the infants in this experiment showed schedule sensitivity on each of these three indices. This *scalloped* pattern of responding to FI schedules and general schedule sensitivity are of course characteristic of rats, pigeons and other animals, but notoriously difficult to replicate with adults and children over the age of four (Lowe et al., 1983).

What are the changes responsible for this transition from animal to human-like responding? Bentall, Lowe and Beasty (1985) addressed this question in a continuation of the previous experiment, but with children of four different age groups; 6 months to 1.5 years, 2.5 to 4 years, 5 to 6.5 years, and 7.5 to 9 years old. Reinforcers of various kinds

were given to the children in accordance with a range of FI schedules which progressed through a range of 10 to 70 seconds for each group. *Low-rate* responding refers to response frequencies which are widely spaced and cumulatively low, while *high-rate* responding refers to response frequencies which are steady, fast, and cumulatively high. Both high and low-rate responding are typical of humans with FI schedules but virtually never seen in animals, which typically produce a more scalloped pattern (Weiner, 1969). The children in the two oldest groups generally showed patterns of schedule insensitivity characteristic of adult human exposure to FI schedules; a failure to show a scalloped pattern of PRP's, and continually demonstrating either high *or* low-rate responding. When the children in these groups were questioned about the contingencies they had faced, those who exhibited high-rate responding reported that they thought there was a relationship between the total number of responses and the likelihood of reinforcement. Those who showed low-rate responding reported that there was a time-related relationship between their responses and reinforcement. In fact, some of the low-rate responders were observed counting up until the time reinforcement was available, enabling them to respond at the appropriate time.

Interestingly, the response patterns of the children in the 2.5 to 4 year old group were much more varied, showing both elements of human *and* animal-like FI responding (e.g., occasional PRP's and periods of high-rate responding), and could not accurately describe the contingencies verbally. The infants' responses were similar to those in the previous experiment, virtually indistinguishable from animal FI patterns. As to be expected, there was a direct correlation between the ages of these children and their performance on the *Reynell Developmental Language Scale* (Reynell, 1977). The authors (Bentall et al., 1985) concluded that the variance across the children aged between 2.5 to 5 years old possibly occurs because of a transition from animal to human-like schedule responding occurring during this period, and that *the development of an ability to formulate rules seems to be at least partly responsible for this transition.*

In other studies with children, a more direct analysis of RGB has been undertaken. For example, Pouthas, Droit, Jacquet and Wearden (1990) employed the verbal report method with 18 children of differing ages in a developmental experiment. The children were either 4.5, 7, or 11 years old. The children were asked to respond when cued by

pressing a button, and reinforcement for "button-pushing" consisted of three levels of feedback. "Very-very-good" was earned by a response latency of 4 to 6 seconds, "very-good" was earned by a response latency between 2.5 and 4 seconds or between 6 and 7 seconds, and "not-very-good" was reserved for responses outside of the 2.5 and 7 second parameters. The children in the oldest group were given this feedback verbally, but for the two youngest groups, the feedback was administered by presenting the child with the face of a clown (e.g., a picture of a laughing clown was a substitute for a very-good score). Each age group was split into an *interview* condition and a *probe* condition. In the probe condition, children were asked after each trial to report how they responded in order get the feedback they were awarded (e.g., "what did you have to do to get very-good?").

The children in the interview condition were not asked about their performance until the experiment was over. The results showed a direct correlation between age and an ability to verbalise contingencies. None of the children in the youngest group were able to say that there was a connection between the response delay and the type of feedback earned, and only one of the children in the middle aged group did so. However, rules describing a direct correlation between latency and the strength of reinforcement were common in the oldest group. Furthermore, children in the probe condition earned a significantly higher proportion of "Very-very-good" reinforcers. Thus, the superior performance of children in the probe condition led Pouthas et al. (1990) to surmise that if *conditions for rule-formulation are optimised, responding is likely to increase in proficiency*. Overall, however, these developmental studies support the hypothesis that the acquisition of *rule-governance is a key feature of human verbal development*.

Employing adults as research subjects has allowed behaviorists to examine the more complex features of RGB. The earliest adult experiments were chiefly concerned with seeking support for RGB as a class of operant behaviour. An example is the landmark experiment by Galizio (1979) which involved exposing four college students to a multiple loss-avoidance schedule. A loss of five cents, indicated by a red light and an audible tone occurred every 10 seconds unless avoided with a response (handle-turn). The programmed schedule changed every 12.5 minutes over each 50 minute trial (i.e., four stages in each trial). The avoidance value of the handle-turn randomly moved

through four stages. In each stage a response would postpone monetary loss for 10, 30, or 60 seconds, and there was an additional stage in which loss was never scheduled. Each of the four schedule conditions were indicated by one of four corresponding amber lights; the subjects were not made aware of this connection, however, and were only told that "a handle turn could prevent a flashing of the red light for 10 seconds or longer". In the first phase, only one out of the four students responded with appropriate sensitivity: that is, at a high rate in the 10 second postponement stage, lower in the 30-second stage, lower still on the 60 second stage, and an almost complete cessation of responding in the no-loss stage. The other three students responded at approximately the same rate throughout the four stages. In the second phase, verbal rules were placed above each of the amber lights so that the subjects could discriminate as to which schedule each of the lights indicated. The labels were simply; 10 SEC; 30 SEC; 60 SEC; and "NO LOSS". Two additional students were introduced at this point.

All six students showed sensitive responding with the labels present. In the third phase, the labels were withdrawn and the amber lights were shuffled. Two of the three subjects who had not shown appropriate discriminations with respect to the amber lights in the initial phase *now began to do so,* even though the signal lights had changed places. Thus, not only had the rules markedly improved responding efficiency, but the generalised behaviour of "rule-following" itself had persisted into a relatively novel situation without the rules, and varied in accordance with the active contingency. Thus, these results would suggest that *rule-following as a class of behaviours can generalise to novel situations and can also appropriately adapt to the contingencies found in novel situations.*

Galizio (1979) also used the same subjects to examine the *extinction* of rule-following. First, appropriate instructional control was re-established using the same labels of 10 SEC; 30 SEC; 60 SEC; and "NO LOSS", again with an amber light indicating which of the four contingencies was currently active. In the second phase, the labels had no correspondence with the actual contingencies indicated by the lights, which in each stage was a no-loss schedule (i.e., a total cessation of responding would still not result in monetary loss). Each of the four subjects still responded at rates which were in compliance with the rules. This was followed by a third phase in which the corresponding

schedules were reactivated, but each response (handle turn) now produced only a 10 second postponement of monetary loss. That is, with the exception of the 10 SEC label, following the rules *would now result in overall monetary loss* because the subjects had direct contact with the programmed penalty. Each of the subjects responding stabilised at approximately 8 to 12 responses per minute, regardless of which of the four contingencies was indicated as current (i.e., ignoring the rules but still responding more often than the required rate of one lever-turn every 10 seconds). In a fourth and final phase the universal no loss contingency was reintroduced (i.e., no loss possible), but contingency-appropriate responding did still not resume, and was almost identical to that exhibited in the third phase. These results would first suggest that if the following of a rule is punished it will mostly likely extinguish and the rule will cease to control behaviour. More importantly, however, this was evidence that *if rule-following does not make contact with the active contingency, and this results in the avoidance of negative consequences, it will be maintained even if it is not the most efficient or appropriate pattern of responding* (i.e., even if it is insensitive). Finally, the results would also suggest that *rule-following which is insensitive will typically persist until immediately followed by negative consequences*. That is, in certain conditions the behaviour of "ignoring rules" can generalise to a novel situation, particularly in the absence of possible punishment.

Later research revealed more about the ways in which rules interact with contingencies. Buskett and Miller (1986) conducted an experiment using college students which studied the interaction of an FI schedule with several types of antecedent rules. At first the subjects were exposed to an FI30 schedule using monetary reinforcement. After a stable baseline was established, the subjects were divided into four groups, one group received no instructions, the other three groups were given the rules that they only needed to respond 15, 30, or 60 seconds after the last reinforcer was earned, respectively.

The results showed that after four trials with the FI30 schedule, PRP's were closest to 30 seconds for the 15 and 30 second groups, but PRP's approximated 60 seconds for the 60 second group. Thus, the 15 second group had strayed from the rule and returned to PRP's approximating 30 seconds in duration, and reported after the experiment that the rule seemed incorrect because the interval was closer to 30 seconds. The 30 second group

followed the rule by approximating 30 second PRP's, and at follow-up interview reported that they thought the rule had been an accurate description of the contingency. However, the 60 second group maintained long PRP's approximating 60 seconds, and failed to deviate from the rule and appropriately adjust responding to the new contingency – the 60 second group even reported that the rules were roughly correct and believed that the intervals ranged from 45 to 60 seconds (note that even though this group only minimally reduced PRP's, nearly every response was immediately reinforced because they still occurred well after the inter-reinforcement interval had expired).

Based on the behaviour of the 15-second group, Buskett and Miller (1986) concluded that when rules conflict with a contingency, and following the rule results in a notable lack of reinforcement, the contingencies may override the rule and eventually regain control of responding. Also, they concluded from the responses of the 60 second group that when rule-following is consistently reinforced it remains strong, even when such compliance fails to produce as much reinforcement as would have been earned if behaviour had been directly shaped by the conflicting contingency (i.e., insensitivity is maintained). To summarise, the results of the Buskett and Miller experiment were similar to Galizio's (1979) findings, and showed that *even when there is a large discrepancy between the rule and the prevailing contingency, rule-following appears to be maintained if it continues to earn reinforcement*. This can occur even if the rule-induced behaviour is not the most efficient or appropriate pattern of responding. However, if *rule-following is reinforced too infrequently, behaviour may abandon the rule and approximate the contingency*.

The Pouthas et al. (1990) study described previously showed that when questions were presented to children that evoked the behaviour of rule-formulation, behaviour increased in efficiency. Thus, one might ask "can rule-formulation itself be taught and shaped like other operant behaviours?" Recall, rule-formulation is precurrent behaviour which is reinforced with the generation of a discriminative stimulus, which then serves to evoke a reinforceable response (some might call the rule a *conditioned reinforcer* from the perspective of the speaker). The rule may not be an *accurate* description of the contingency (in fact it may not even be a contingency), but from the *speakers* perspective, a contingency has been described and this statement can be used to evoke

RGB. Catania, Matthews, and Shimoff (1982) conducted an experiment with 50 undergraduate subjects, each concurrently exposed to both random interval (RI) and random ratio (RR) schedules. An RI interval constantly varies the period for which responding must have ceased before a reinforcer is available based on an average parameter, and thus responding relatively slowly is the most efficient way to obtain maximum reinforcement (i.e., maintaining relatively *low-rate* responding is optimal). An RR schedule constantly alters the total number of responses required before the administration of a reinforcer based on an average parameter, and the faster one responds the more reinforcement will be earned (i.e., consistent *high-rate* responding is optimal). One of these two schedules was always in operation during each trial, as indicated by a corresponding light on the left or right of the stimulus panel. The two schedules would successively alternate as active, changing every 1.5 minutes. Occasionally, the RR and RI schedules would reverse positions also (i.e., sides). The subjects could earn points which could later be exchanged for money by responding with button presses, but could also earn points by trying to formulate rules about the pattern of responding required by each schedule. This "guessing" stage occurred between trials for a duration of two minutes, and required the subject to write down all of the possible ways in which responding could earn reinforcement for each of the two schedules. The examiner would then award 0, 1, 2, or 3, points for each rule depending on its level of correspondence with the RI or RR schedules (i.e., how correct it was). For example, a rule which simply specified that "speed of pressing" was the critical factor in earning reinforcement from the RR schedule would earn three points. Other rules which only contained a reference to topography (e.g., "you just need to push the button hard") or inaccurate rate patterns (e.g., "I had to push the button slow, then fast, then slow again") earned fewer points. Additionally, a subgroup of the subjects had their rule-formulation behaviour *shaped* by being immediately informed of the points that each constructed rule had earned. Later in the experiment, another separate subgroup was actually *instructed* on how to formulate rules about the schedules (i.e., *rule-governed* rule-formulation). For example, the instructions were given to "write 'press slowly' for the right-hand button". Typically, alternating RR and RI contingencies will produce a correlation between response rates required and differential response rates emitted when presented in single succession. RR will increase

response rates, RI will slow responding down (Matthews, Shimoff, Catania, & Sagvolden, 1977). Differential rates of responding which were in accordance with the active schedule were used as a measure of non-verbal sensitivity (i.e., faster for the RR and slower for the RI would represent sensitivity). The extent to which the generated rules differentiated between the RI and RR schedules was used as a measure of verbal sensitivity (i.e., the frequency of the rules suggesting slower responding for the RI schedule and faster responding for the RR schedule). The extent to which the subjects behaviour corresponded with the rules they had generated was used to estimate how influential the rule was.

The results showed no relationship between verbal and non-verbal sensitivity (i.e., verbal behaviour sometimes controlled, was sometimes controlled by, and was sometimes independent of non-verbal behaviour). However, the rules constructed by the subjects who had their rule-formulation shaped almost always followed those rules accurately, even when the rules were inconsistent with the active contingency. However, when the shaped-rule subjects earned minimal points for rule-formulation or rule-formulation was not differentially reinforced (i.e., when they provided poor rules or when there was no variance in rule-quality), their responding was rarely in conformity with the rules. The subjects who were simply told *how* to formulate rules did so with accuracy, but varied greatly as to whether they actually *followed* those rules. In summary, the results suggest that *although rule-production may require prolonged periods of successful shaping before it effectively controls behaviour, once established, those rules control behaviour more reliably than rules which have been simply presented to the subject outright.*

We know that rules control behaviour and can even produce insensitivity to prevailing contingencies, but little has been said about the nature of the formal properties of those rules. Can certain types of rules lead to certain types of responses? Matthews, Catania, and Shimoff (1985) conducted a study which examined the effects of two fundamental types of rules. Seven undergraduate students were exposed to multiple random ratio (RR) and random interval (RI) schedules as in the previous experiment outlined. A green light would indicate that reinforcement was available for a response. As before, subjects could earn points which could later be exchanged for money in two ways: by responding appropriately to either of the two schedules, or by formulating rules

about the randomly altering schedules. However, this time the subjects could only write three rules for each of the two schedules. Four of the students were awarded points based on "performance" descriptions, and the other three students earned points for general "contingency statements". Rules were assessed by using two types of rule-questionnaire sheets for each group. The contingency statement sheets required the subject to complete the rule-statement "The computer will let your press turn on the green lights depending on...". Alternatively, the performance rule sheets required the subject to complete the rule-statement "The way to turn the green lights on is to...". For example, in the performance-description group, a rule that described the RR schedule of reinforcement as being available through "a changing number of button presses" would receive fewer points than a rule specifying that the way to attain reinforcement from the RR schedule was "to press the button as fast as possible". The system of scoring used accounted for many rule characteristics, and proportionally dictated points (e.g., reference to bodily movement always earned three points for performance-related shaping). The results were similar to the previous experiment by the same authors (Catania et al., 1982), and showed that shaping was a more important determinant than the schedules, but another interesting pattern emerged. The performance-statement group produced a significantly higher correlation between rules and induced compliance. In some cases, even when the RR and RI contingencies reversed positions, performance-rules maintained responding as though no change had been made. For the contingency-statement shaped group, the correspondence between rules and pressing rates were much more variable and largely unrelated. Based on these results, the authors suggested that the history of reinforcement humans have for rule-following may lead to *performance-specific rules being stronger determinants of behaviour than general "pattern related" rules or contingency statements.*

All of these studies suggest that a typical product of RGB is behaviour which is in some way insensitive to the prevailing contingency. This is probably not surprising if one accepts that RGB is a general response class that is as highly strengthened and ubiquitous as Skinner (1969) argued. However, it should be noted that as a behavioural phenomenon, rule-induced insensitivity depends on many variables, and thus one should not make a complete generalisation of these findings to circumstances outside of the

laboratory. Furthermore, the results of a more recent study by Newman, Buffington, and Hemmes (1995) has shown that RGB can be as much a function of the schedule of reinforcement that reinforces compliance as the nature or accuracy of the rule itself. It is also important to note that many other interesting studies on RGB have also been conducted in the last 10 years or so, and the studies I have discussed are only the most basic examples of an enormous body of literature.

At the very least, however, the empirical research on RGB would suggest that rule-making and rule-following can be meaningfully spoken about as operant behaviour, and may play an important role in human development. It would also appear that rules can produce insensitivity to prevailing contingencies unless there are immediate and aversive consequences for doing so, and shaped self-rules exert more control over behaviour, especially when such self-rules are performance-specific. These claims are not hard to make sense of. People do tell themselves what to do, and this presumably on occasions allows one to maintain responding in the absence of reinforcement. Furthermore, it does often appear as though humans are reluctant to change the rules they are following until there is an apparent correspondence between their rule-governed actions and the negative consequences they produce. Picture the person who uses a map to find her way home from work, but rarely takes the shortest path. Solving the problem of finding a quicker way home may not be a priority until it is found that the first 10 minutes of a rescheduled television program will be missed unless a shorter route is devised (i.e., a new rule). Furthermore, holding all other factors constant (e.g., the authority of the speaker), it is often the case that we put faith in the conclusions we have come to ourselves far more than those told to us by others. This observation undoubtedly alludes to the finding that shaped self-rules can control behaviour more reliably than externally-provided rules. Finally, if we can be relatively specific about our plans and the actions they detail, they are usually a lot more influential than general ideas or broad descriptions of the situation at hand. Thus, performance-specific rules may simply seem easier to follow. In summary, Skinner's (1969) claims about the ubiquity of RGB seem to have come to life in applied behavioural research. As I will discuss next, however, verbal behaviour and rule-governance are also related to a much bigger picture; the development of the verbal repertoire and the evolution of human culture itself.

1.6 RGB and the Law-of Effect.

Earlier I gave a brief description of the way a functionalist perspective looks at operant and classical conditioning when evaluating the interaction between environment and organism. What I purposely omitted then was the fact that functionalism is closely linked to the theory of biological evolution proposed by Charles Darwin (1809-1882). Only now, having discussed the intricacies of RGB, are we ready to place the theory within the broader context of biological determinism.

Darwin proposed that a species' physiological structure or *genotype* is the result of thousands or even millions of years of differential success across the biological variations of a species, where *success* is defined by an increased capacity to reproduce. This type of evolution or *phylogenic* evolution requires three things to occur. There must be a limited amount of resources for which the members of a species are competing. That is, *competition* will allow certain physical characteristics to determine which are the most useful in obtaining those resources which contribute to increased success. There must also be *variation* across members if some characteristics are to prove more advantageous than others. Finally, a species must *reproduce* regularly; without reproduction there would be no way for the members with the more successful characteristics to transmit these advantages to their offspring.

Operant behaviour evolves throughout one's life in a manner analogous to that of phylogenic evolution. This process is commonly referred to as *ontogenic* evolution (Baum, 1994). Ontogenic evolution requires the same three ingredients phylogenic evolution does. There must be *competition* between behaviours, in that the environment selects behaviours by rewarding some with more or less reinforcement than other topographies. There must be *variation* across behaviours, so that the operating contingencies can differentially reinforce and select the most effective topographies. There must also be *reproduction,* or repeated trials of the behaviour, so that shaping may occur over successive instances. Just as phylogenic evolution increases the suitability of a species' genotype to its environment, ontogenic evolution increases the suitability of a person's behavioural repertoire to her environment. Phylogenic and ontogenic evolution share a common basis with their relationship to the *law of effect,* a term coined by

Edward Lee Thorndike (1874-1949) to describe the ubiquitous tendency for natural phenomena to strengthen when followed by a *satisfier* or a *pleasant event*. Applied to behaviour, the law of effect was employed in Skinner's work as it provided a scientific account of organismic activity which did not necessitate a reliance on mentalistic explanations. As verbal behaviour is operant behaviour, Skinner was adamant that the concept of ontogenic evolution could be extrapolated to explain the development of the verbal repertoire in the individual (Skinner, 1986).

Every verbally-capable individual undergoes an ontogenic verbal evolution over the span of a lifetime. The limited repertoire of vocal sounds that we are endowed with at birth are reinforced and shaped by others, and we soon learn a range of actions, each of which produce verbally-mediated consequences. Before long we have learned to respond to the many different sounds and signs produced by others, and eventually recognise statements that *specify* a set pattern of responding (i.e., rules). Our first contact with rules is probably provided by our parents. Although we may have not always followed the rules we were given when young, we would have responded in some way, allowing our parents to differentially reinforce our varying approximations of rule-following. We come to respond to many variations of rules as a class of stimuli, possibly described as those indicating "an availability of reinforcement for compliance". People follow thousands of rules produced by themselves and others throughout their lives (variation), and through repeated instances of rule-following (reproduction), find that some are more effective than others in obtaining reinforcement (competition). Thus, at the level of the individual, RGB evolves by the same ontogenic process of shaping that many non-verbal behaviours do.

On a wider scale, the concept of RGB can be used to understand the evolution and maintenance of cultures in terms of the contingencies acting on its members (Baum, 1995; Malott, 1988). It is clear that a culture's ability to learn from its elders is undoubtedly an invaluable asset to the survival of its members. Perhaps the advantage of being able to verbally transfer stimulus control from generation to generation is most commonly evidenced by the construction and presentation of rules which hold their applicability over time. For example, if every generation needed to have experience with incest in order to learn about the aversive effects of inbreeding, many individuals would

be affected with serious genetic defects. Accordingly, a maladaptive behaviour such as incest runs counter to the rules of most cultures (Skinner, 1974). Thus, a verbal community can avoid the need for each member to have had direct experience with a contingency in order to emit effective responding, if directions for that contingency can be verbally learned and taught.

If each generation had to go through the slow and awkward process of formulating its own basic societal rules, humans would probably never develop rules beyond those based on the contingencies experienced in the lifetime of one generation. Over many generations, a culture experiments with a range of different cultural rules (variation), and is successively faced with many circumstances which require its members to follow that practice (reproduction). Through repeated attempts at rule making, a culture learns to find that some rules are more effective than others in obtaining reinforcers that benefit the group as a whole (competition amongst rules). Thus, the repertoire of cultural rules that any verbal community might learn necessarily evolves over time. Skinner (1986) frequently spoke of the acquisition of verbal behaviour as marking a crucial turning point in the acculturation of human beings. Finally, as Baum (1995) rightly points out, unless the concept of RGB is seen as a derivative of evolutionary theory, it risks becoming detached from the affiliation with biology and environmental determinism that other areas of behavioural science continue to enjoy.

1.7 RGB as Applied to Cognitive Psychotherapy

What are some of the direct implications of radical behaviourism's advancements in the analysis of RGB in relation to contemporary psychotherapy? First, the importance of the concept of RGB becomes most apparent when one considers the fact that behaviour analysis has received much criticism for a supposed *inability* to deal with fundamental human processes such as thinking, memory, problem-solving, imagery and language (Marr, 1984). Much of this may stem from confusion between the philosophies of *methodological* behaviourism and radical behaviourism. Because methodological behaviourism is commonly used as a synonym for *empiricism*, the current emphasis on

empirical research may have denigrated the importance radical behaviourism places on private events (Zettle & Hayes, 1982). However, it is equally likely that such criticisms have partly arisen out of concerns about the limitations of generalising principles of basic animal behaviour to humans (Galizio, 1987). For example, basic laboratory studies clearly show that unlike animals, humans often fail to make appropriate changes in responding when schedule changes are made (Galizio, 1987). Nevertheless, behavioural techniques *derived* from basic learning principles are clearly effective in treating a wide range of psychological maladies, including major depression, anxiety disorders, addiction, sexual disorders, marital and family problems, and many others (for a review of the effectiveness and scope of behaviour therapy, see Emmelkamp, 1994). However, the disconnection between applied and basic behaviour analysis is a cause for concern. One obvious problem is the apparent inapplicability of an enormous quantity of animal behaviour research to the understanding of human behaviour (Galizio, 1987; Mace, 1994).

It is possible that these criticisms of applied behaviour analysis have contributed to a gradual immersion of cognitive methodology into what were once purely behavioural treatment programs (Mace, 1994). I will fully elaborate on the nature of the cognitive paradigm in Part Two, but for the present purpose cognitive therapy can be defined as "the application of the cognitive model of a particular disorder with the use of a variety of techniques designed to modify the dysfunctional beliefs and faulty information-processing characteristics of each disorder" (Beck, 1993, p. 194). Cognitive approaches have been effectively applied to the treatment of many types of mental disorders and psychological distress (see Hollon & Beck for a review of the efficacy of cognitive therapy, 1994). However, probably the most striking advancement in psychotherapy over the last two decades has been the development of cognitive therapies and their combined use with behaviour therapy. Treatment packages combining elements drawn from the two schools have come to be called cognitive-behavioural therapy (CBT). Currently, 55% of all clinical practitioners in New Zealand refer to themselves as predominantly cognitive-behavioural (Kazantzis & Deane, 1998). Research conducted in the United Kingdom and the United States show a similar popularity of this integrated approach to therapy (Jensen, Bergin, & Greaves, 1990).

Empirically speaking, the impetus for combining cognitive and behavioural therapies is unclear. A meta-analysis of 29 outcome studies conducted between 1970 and 1989 tested purely behavioural therapies with those that added cognitive therapy to those same methods; 83% reported that the addition of cognitive components resulted in equivalent or inferior outcome (Sweet & Loizeaux, 1991; Latimer & Sweet, 1984). However, of equal importance is the fact that as a unitary field of psychological investigation and practice, the cognitive-behavioural paradigm has experienced great difficulty in trying to establish a firm conceptual basis (Mahoney, 1974; Zettle & Hayes, 1982). First, differing truth criteria put the philosophical stances of the two schools at odds. While the behavioural model excludes any notion that refers to events or matter that do not exist in space and time, the cognitive model allows and often necessitates the use of dimensionally fictional entities (i.e., hypothetical constructs) in explaining behaviour. Second, notions of causality differ greatly. Many cognitive models rely on the use of information-processing models to map thought processes as they relate to overt activity (e.g., analogies with computer metaphors), while behaviourism sees the ultimate causes of such activity as purely environmental, serving the evolutionary function of increasing an organism's suitability to its surroundings. Given such obstacles, it seems unreasonable to expect any attempt to integrate cognitive and behavioural theory to produce much more than an uncomfortable amalgamation of fundamentally disparate concepts.

Such controversy has led some to speculate on the utility of the concept of RGB as an integrative theme for CBT. Because cognitive therapy is primarily verbal or semantically oriented, a functional analysis of cognitive therapy in terms of RGB might assist in our efforts to bring the two schools under a unified theoretical base. In their article outlining a proposal for such an integration, Zettle and Hayes (1982) state:

> It is our view that the concept of rule-governed behaviour, if expanded somewhat, can incorporate much or all of cognitive and semantic therapy without distorting its basic phenomena. It also points to several mechanisms of change as yet under-investigated in the field (p. 76).

Similar ideas have been expounded on by other researchers – the argument that the concept of RGB could be used as a way for behaviourism to "claim" some of the

strategies and techniques used in cognitive clinical interventions by thoroughly explaining these activities within its own behavioural paradigm (e.g., Jaremko, 1987; Poppen, 1989). A major implication of this, and an objective anticipated by the present essay, is that a theoretical integration of cognitive and behavioural therapy may point behaviourists (and cognitivists for that matter) to *aspects of human behaviour that have previously been ignored.* Such awareness may result in both scientific and applied advantages. Ultimately, however, the driving force behind this type of integration lies in the assumption that a paradigm holding some theoretical universality and freedom from conflicting fundamentals may both develop and expand more rapidly (Arkowitz, 1989). With this, I will now turn to Part Two and outline four of the most influential cognitive therapies and provide a more elaborated discussion on what it might mean to use Skinner's theory of RGB to integrate the cognitive and behavioural therapies.

PART TWO

COGNITIVE THERAPY AND PSYCHOTHERAPY INTEGRATION

A failure on the behalf of any integrative project to fully outline each of the competing theories involved is invariably at the expense of appearing sectarian. Part Two is primarily devoted to an introduction of the major themes of four highly influential cognitive approaches, the core points made by their leading proponents, and their current empirical status. Specifically, I will describe the *cognitive therapy* (CT) of Aaron T. Beck (1967; 1976; 1993), the *rational-emotive behaviour therapy* (REBT) of Albert Ellis (1962; 1995a; 1995b), the *self-instructional therapy* (SIT) approach as espoused by Donald Meichenbaum (1977; 1985), and finally, the *social-cognitive* or *self-efficacy* theories of Albert Bandura (1977a; 1977b; 1986). At that point, having outlined both the behaviourist's position on verbal behaviour and several forms of cognitive therapy, I will then define what it means to attempt a rapprochement of these two fields. I will provide an introduction to the history and aims of the field of psychotherapy integration, and present the many forms that integrative work may take. Finally, I will detail and justify a selection of one type of integration to be employed here: *assimilative integration* (Lazarus & Messer, 1991). By then I will have laid an adequate grounding for Part Three, which demonstrates an application of the concept of RGB to an integration of cognitive and behavioural therapy.

2.1 The Cognitive Paradigm.

As a philosophy, the cognitive approach has been defined as the school of thought concerned with "all processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used" (Neisser, 1967, p. 24). It is difficult to precisely trace the beginning of the cognitive movement. In many respects, the psychoanalytic school of Sigmund Freud (1856-1939) was cognitive in that it relied on inferred mental constructs which could not be directly observed. However, some would argue that Watsonian theory played as much a part in the events that gave rise to cognitive philosophy as any other tradition (Reed, 1992).

According to Beck and Weishaar (1995), the first piece of work to be closely aligned with contemporary cognitive psychology was probably that produced by George Kelly (1905-1966). Kelly's (1955) book "The Psychology of Personal Constructs" began to look at the role of belief systems in terms of hierarchical levels of cognitive organisation, which was in great contrast to the behavioural and psychoanalytic zeitgeists of that period. However, Kelly alone was by no means responsible for the emergence of the cognitive paradigm, as his work is merely reflective of the underpinnings of the tradition. A factor more commonly cited as responsible for the turn towards cognition is the gradual recognition around the early 1960's of several anomalies in purely behavioural approaches:

> The development of the cognitive sciences and the re-emergence of cognition as a legitimate scientific pursuit within psychology symbolised an excursion into what the behaviourists termed the *black box*, this time with novel approaches and new tools to study its content, organisation and processes. This renewed interest in cognition served to fuel a trend in behavioural psychology generally characterised as a shift from a behavioural/associationistic to a cognitive/mediational perspective. Following Breger and McGaugh's (1965) famous "reformulation and critique" of learning theory, research evidence contrary to the predictions of traditional conditioning theory became increasingly visible. The data made it clear that non-mediational accounts of human learning could not accommodate the complexity encountered in experiments and therapeutic interventions with humans. (Mahoney & Lyddon, 1988, p195).

However, the emergence of the cognitive paradigm can not be solely attributed to deficits inherent in rival traditions[7]. Psychologists were genuinely interested in the cognitive approach, and the execution of successful research within the school ensured that its popularity grew (Mahoney & Lyddon, 1988). An alternative is to suggest that unlike B.F Skinner's claim to radical behaviourism, or Ivan Pavlov's claim to the

[7] It should be noted that several theorists *do* argue that the emergence of the cognitive paradigm occurred primarily as a result of increasing criticisms of behaviourism (see Dember, 1974).

classical conditioning paradigm, the cognitive school was not a result of any one discovery or leading proponent. Instead, the emergence of cognitive psychology seemed to be more reflective of a *general paradigmatic shift* involving a number of scholars taking an entirely new approach to the study of behaviour (Reed, 1992).

With the exception of structuralist cognitive psychology, best represented by the works of developmental psychologist Jean Piaget (1896-1980) and his followers, the bulk of cognitive theory is based on the concept of *information-processing*. Information-processing attempts to analyse the ways in which data are received, encoded and then prepared so as to be emitted in the form of a response. These analyses often take the form of elaborate models which detail the separate stages through which information is said to pass. Below is a diagram typical of those found in information-processing models, taken from Reed (1992)(Fig. 2.1). The double-ended arrows indicate that information can flow in both directions, and if necessary, can be reprocessed many times. *Sensory store* is raw information directly obtained from the perceptual senses. This information is then *filtered*; data that are recognised as forming a concept or pattern are retained while redundant sensory information is discarded. *Pattern recognition* can involve any organised referent or item that is familiar to the person (e.g., a car, horse, etc.). Of these patterns, some are *selected* and stored in short term memory (STM), often so as to be manipulated for the design of a response (i.e., working memory). Other information may be stored in long-term memory (LTM) for later use.

Figure 2.1 A typical information-processing diagram.

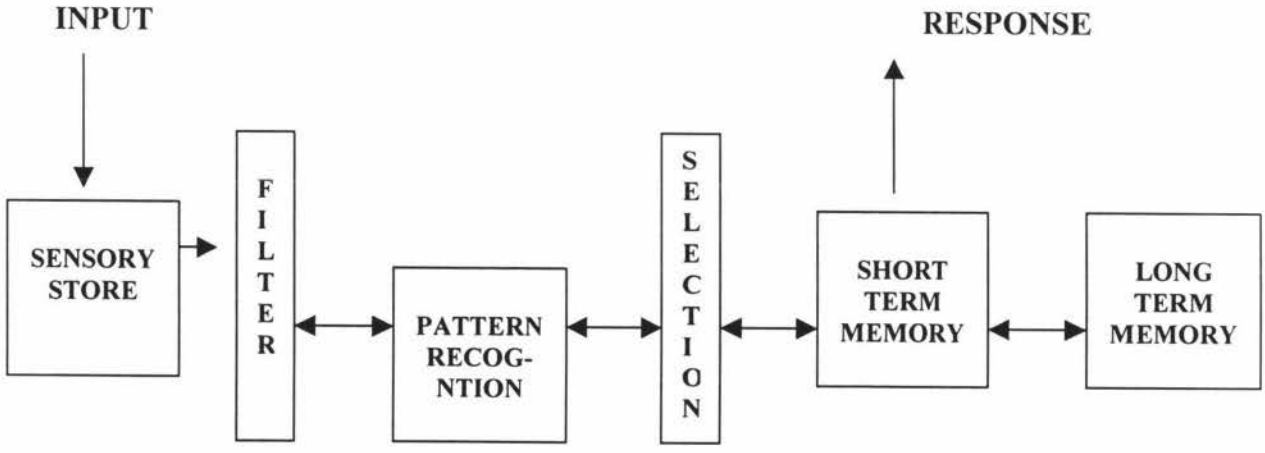


Figure 2.1 is a simplified model, and diagrams tracing decision making, problem-solving, and other higher executive functions are often much more complex (see Benjafield, 1997). The point is that in contrast to the radical behavioural approach of reducing discrete mental events to their essential measurable features before examining them, the cognitive model disposes of such a need and tries to diagrammatically organise the events that occur *between* an overt stimulus and its response. Reese (1989) points to the Turing-Church thesis as having provided a fundamental rationale for the information-processing approach. A Turing machine is a hypothetical computer that can calculate anything calculable by machine, no matter how complex. The Turing-Church thesis holds that *if a problem that could be presented to a Turing machine is not solvable by a Turing machine, then it is also not solvable by human thought – human thought can therefore be adequately modelled by a machine* (Kurzweil, 1985). Although Reese has identified several logical errors associated with this postulate, the Turing-Church thesis is a classic example of the computer metaphors that characterise cognitive psychology. Perhaps, then, cognitive psychology could probably be said to rely as much on logical inference as it does on observation.

So far I have only mentioned some of the most abstract fundamentals of the cognitive paradigm. Because cognition can be analysed as fundamental human phenomenon, human dysfunction can also be conceptualised at the cognitive level, and similarly, rectification of this dysfunction can be approached by using the tools of the cognitive psychologist to encourage adaptive change. Thus, the utility of the cognitive approach to psychology is perhaps best illustrated by the vast range of therapeutic strategies that are based on the cognitive model of behaviour. I will now proceed to outline four different models of this orientation.

2.2 The Cognitive Therapy of Aaron T. Beck

Aaron Beck (1967, 1976, 1993) is most commonly referred to as the founder of Cognitive Therapy (CT), possibly because of the profound impact his early research on depression made (see Beck, 1967). In fact, the term *Beck's CT* is often used

synonymously with reference to "general CT". Here I will use the terms interchangeably, but only when referring to other forms of CT which are closely aligned with Beck's.

CT can be thought of as consisting three main parts: (a) a *core theory*, (b) a system of *strategies*, and (c) a series of *techniques* (Beck and Weishaar, 1995). The core theory of Beck's approach holds that many psychological disorders stem from a *systematic bias* in information-processing. A *bias* can be defined in the current context as a distortion in the interpretation of certain events. The nature of the bias typically determines the types of disorder or problems that the patient suffers. For example, manic symptoms are often characterised by an exaggerated perception of potential gain, self-worth, and personal capability. According to Beck (1976) these distortions normally take seven major forms:

Arbitrary Inference: A conclusion that has been arrived at in the face of conflicting data.

Selective Abstraction: A conclusion or conceptualisation of a circumstance based on a specific isolated detail, ignoring the context from which that information was taken.

Overgeneralisation: Using a context-specific conclusion to make general assumptions about a wide range of circumstances, many of which may be totally unrelated.

Magnification and Minimisation: Interpreting an event as being far *more* significant or far *less* significant than it actually is.

Personalisation: Interpreting an event as having direct relevance to oneself without clear support for a causal connection.

Dichotomous thinking: Categorising events as either complete success or total failure, rarely identifying with the mid-ground between these two extremes.

Ignoring the Positive: Selectively attending to negative events and thoughts while ignoring the positive aspects of one's surroundings.

Not only does Beck focus on the specific *content* of biased processing, but the underlying cognitive sets which give rise to these distortions are elucidated. These cognitive sets are more often referred to as *schema*. Schema are sets of attitudes which serve to organise one's way of thinking about the world and its contents. A schema is comprised of very general beliefs which allow us to categorise the meanings we ascribe

to events. For example, the belief that "a university degree leads to a higher social status" might be part of a firmly held schema about education. Schema contribute to biased processing because they typically serve as an individual's frame of reference when deciding on how to interpret a situation or event.

Beck's CT aims to challenge biased information-processing in several ways. Interview techniques rely heavily on the *Socratic dialogue*. The term is derived from the method of dialogue used by the Greek philosopher Socrates (470-399 BC). The Socratic dialogue is a method of discussion which involves using nothing other than directive questioning to evoke and convey information. Through *guided discovery* the client comes to realise that many past situations elicited can be seen as a result of similar cognitive distortions, and that new experiences, possibly in the initial form of homework exercises, can allow the practice of more adaptive perspectives. *Collaborative empiricism* involves the patient and therapist exploring these logical distortions in a co-operative and scientific way, within a supportive team-work atmosphere.

The first goal in CT is to identify the *automatic thoughts* which typically follow upsetting events. By examining the available evidence, the client and therapist can evaluate how accurate such thoughts are, and what distortions may be present. For example, a client may be encouraged to adopt the habit of writing down the situation responsible for the emotional upset and the meaning they have attributed to it. Keeping a daily record of dysfunctional thoughts may be used to clarify and simplify these re-evaluations. This process often includes experimental tests to assess whether the client's original inferences were realistic. For example, a client may have to ask somebody what it is they were meaning before assuming a negative message was implied. Once it is clear that maladaptive assumptions have been made, more plausible and emotionally corrective interpretations can be realised. Long-term change occurs when the client independently begins to re-evaluate day-to-day situations in a more adaptive and balanced way, a gradual process termed by Beck (1976) as the *cognitive shift*.

CT is structured, problem-oriented, therapist-directed, and usually short-term, with treatment duration commonly falling within the range of 12-16 sessions (Beck & Weishaar, 1995). CT regards current and ongoing thought processes as the primary focus in therapy, and unlike psychoanalytic approaches, does not place importance on

discovering dysfunctional libidinal urges or other repressed wishes. CT almost always incorporates techniques which have a distinctly behavioural theme also. For example, Beck may encourage clients to seek out new situations and try new activities when testing their hypotheses about the world. CT may also encourage people to evaluate their environments, how they have constructed these environments, and the impact their surroundings have on maintaining current belief patterns. However, unlike behaviourism, which regards the environment as the primary controller of behaviour, Beck argues that the environment is secondary to the influence that interpretation and belief systems can have on our decisions and actions.

Worth special mention here is Beck's widely used and researched CT for depression (Beck, 1967). Beck holds that depressive states are characterised by the *cognitive triad:* a tendency to think negatively about the world, other people, and oneself (Beck & Weishaar, 1995). Because the negative triad pervades all areas of life, depressed clients interpret their situation as both inordinately demanding and inescapable. Subsequently, a sense of hopelessness can lead to a *paralysis of will*, a reluctance to take part in normal daily activities. CT helps the client reveal those logical distortions (as outlined earlier) which may be contributing to the negative triad, and trains the individual to evaluate these situations realistically. Meta-analysis has shown that cognitive therapy for depression is at least as effective as other approaches, including pharmacotherapy (Dobson, 1989).

Studies testing cognitive treatments similar in philosophy to Beck's general approach have been found to be effective in the treatment of bulimia nervosa (Wilson & Fairburn, 1993), more effective than applied relaxation, imipramine, and placebo in the treatment of panic and agoraphobia (Clark, Salkovskis, Hackmann, Middleton, & Gelder, 1994), and more effective than wait-list control and graded exposure with relaxation in the treatment of generalised anxiety disorder (Butler, Fennell, Robson, & Gelder, 1991). Furthermore, when combined with guided exposure, CT is as effective as phenelzine and superior to supportive group therapy and pill placebo controls in the treatment of social phobia (Heimberg & Lebowitz, 1992). Finally, CT can be effective in reducing the amount of conviction held in delusional beliefs (Chadwick & Lowe, 1990). CT has also

been effectively applied to the treatment of many other psychological problems (Hollon & Beck, 1994).

2.3 The Rational-Emotive Behaviour Therapy of Albert E. Ellis.

Independent of Beck, Albert Ellis (1962, 1995a, 1995b) has developed a cognitive therapy which he has labelled rational-emotive therapy (RET). The central premise underlying RET is the claim that most psychological human distress is in some way caused by irrational beliefs. According to RET, when an emotional experience, C, immediately follows an activating event, A, A may be assumed to be the cause of C. However, C may be largely caused by the individuals beliefs about A – this is point B. In the present sense, a belief can be defined as a notion held to be *real* or *true*. Ellis' claims that when C is caused by irrational beliefs they must be disputed (at point D) if they are to be replaced with less emotionally-charged beliefs. Like Beck, Ellis recognises the necessity of behavioural strategies, and thus homework, graded task assignments, and exposure are all commonly used in RET. Because of this, Ellis (1995a) later changed the name of RET to *rational-emotive behaviour therapy* (REBT). The main points of REBT are follows (Ellis, 1995b):

1. Humans have an innate capacity to be rational or *self-constructive* and all ultimately aim to reach their potential. Although we try to think logically in the name of self-preservation and efficiency, we often abandon this capacity for various reasons.

2. Our tendency to think irrationally can be greatly influenced by our culture, and we may be most susceptible to these influences when we are young.

3. Every person is a complex system, acting emotionally, behaviourally, and cognitively. These processes are determined by our history and rarely occur in isolation. Thus, REBT aims to use a variety of perceptual-cognitive, emotive-cathartic, and behavioural-educational methods in order to attack irrational beliefs from multiple directions.

4. Therapies which do not endorse a highly directive and problem-oriented approach as used in REBT are not likely to be as effective in treating most psychological problems.

5. Although REBT stresses that the therapist display unconditional acceptance and close collaboration with clients, the therapeutic relationship is by no means central to the approach.

6. REBT incorporates a range of behavioural techniques outside of the semantic components of therapy. This involves role-plays, desensitisation, operant conditioning, problem-solving skills, and so forth. However, the *learning component of these strategies are secondary to their effects on eliminating irrational beliefs*.

7. Although REBT acknowledges that biological and traumatic dispositions can contribute greatly to psychopathology, the *effects* of these dispositions can almost always be conceptualised as faulty "point B" processes.

8. *Insight* is not of primary importance to REBT in the traditional sense. The only insight of importance to REBT are (1) becoming aware that events, while distressing, often only acquire their painful effects because of the belief systems they operate on, (2) irrational beliefs are systematic and can be so automatic and ingrained that they can subtly pervade all areas of life, and (3) only hard work and a willingness to test the rationality of one's beliefs will rectify this dysfunctional pattern.

9. Stimulus-response theories are undermined by overlooking the sequences of events that occur *between* the stimulus and response.

REBT aims to help people become aware of the unrealistic demands that they make of themselves and others. Ellis (1995b) warns against therapies which try to meet the demands of clients rather than change them. For example, the belief "my therapist must

listen to *all* of the problems I have had in my life before she can accept me" is easily reinforced by the therapist listening to deep-seated emotional problems, and may be a tempting option to the inexperienced clinician. Ellis believes that such indulgence will not contribute to change. Instead, beliefs at point B need to be changed from "absolutistic" to "preferential". Rigid demands involving *should's* (e.g., "I should have a better job!"), and *must's* (e.g., "I *must* get that promotion!") need to become more flexible (e.g., "I would *prefer* a better job", or "I would *like* to get that promotion").

Silverman, McCarthy and McGovern (1992) reviewed 89 psychotherapy outcome studies conducted between 1982 and 1989 that assessed the effectiveness of REBT across a wide range of clinical disorders. These included outcome studies testing REBT alone, testing REBT against other types of treatment, and studies that combined REBT with other approaches. Of the 89 studies, 49 indicated clinically significant results at follow-up. More recently, an applied REBT guide, "Using REBT with common psychological problems: A therapist's casebook", has been published (Yankura & Dryden, 1997). This casebook, supported with detailed descriptions of many single case treatments, explains how contemporary REBT can be applied to obsessive compulsive disorder (OCD), generalised anxiety disorder (GAD), anger related problems, panic disorder and agoraphobia, ADHD, depression and other mental disorders.

2.4 The Self-Instructional Therapy approach of Donald Meichenbaum.

Donald Meichenbaum (1977, 1985) has also developed a method of cognitive therapy, which he has labelled *self-instructional training*. Self-instructional training essentially involves the replacement of negative self-statements with verbalisations that are more adaptive and assist in problem-solving and coping. Self-instructional methods are perhaps best known for their use in *stress inoculation training*. Stress inoculation training, a subsidiary of SIT developed by Meichenbaum (1985), involves training the client to use verbal instruction and self-management to progress through a program of increasingly stressful assignments. However, the literature commonly refers to Meichenbaum's general approach as Self-Instructional Therapy (SIT) as this includes both self-instructional training and stress inoculation training, and other derivative

approaches which incorporate the same philosophy, such as problem-solving training (D'Zurrilla & Goldfried, 1971). Thus, although many self-instructional training programs are multifaceted treatment programs in which self-instruction plays a key role, SIT constitutes a therapeutic method in its own right (Hollon & Beck, 1994)

SIT approaches are based on the notion that self-talk can be used to help a client effectively cope with difficult situations as a *primary* method of change, because self-talk defines both the meaning we ascribe to events and the actions we take in response to them. By reciting directive statements before a task the client is trained to approach a task in a systematic manner so that an effective level of performance is attainable. For example, reciting confidence-related statements before performing a speech such as "I know I can do this because I have done it many times before" would be an example of using self-instruction to deal with an otherwise unbearable task. The client is typically instructed to engage in self-reinforcement immediately following such self-instructed behaviour (e.g., "I knew I could do a great speech"). Thus, the major difference between SIT and other cognitive methods is that the primary emphasis is on *cognitively coping with certain environments rather than altering one's perception of them.*

Meichenbaum's (1985) stress inoculation training is a technique subsumed beneath the wider category of SIT. In some ways it can be considered to be the behavioural paradigm's equivalent of graded task-scheduling, because stress inoculation involves reaching an ultimate goal by progressing through smaller tasks of increasing difficulty. Like SIT, however, the core tenet underlying stress inoculation training is that many psychological problems can be conceptualised as a function of stressful events, and that a *systematic* approach to coping, using internal dialogue and behavioural exercises, can allow one to gain increased control over these events:

> In some ways stress inoculation training is analogous to the concepts of medical inoculation against biological diseases and attitude-change immunisation. Analogous to medical inoculation, stress inoculation training is designed to build "psychological antibodies" or coping skills, and to enhance resistance through exposure to stimuli that are strong enough to arouse defences without being so powerful so as to overcome them. In this way, the client can (a) develop a sense of

> "learned resourcefulness" through experiencing success in coping with manageable levels of stress, and (b) build a prospective defence composed of skills and positive expectations that will help him or her deal effectively with even more stressful situations (Meichenbaum, 1985, p. 21).

Meichenbaum (1985) states that the core aims of stress inoculation training are to:

1. Teach clients the reciprocal and complex mechanisms of stress and coping.
2. Train clients to self-monitor maladaptive thoughts, images, feelings and behaviours so that they may be adjusted accordingly.
3. Train clients in effective problem-solving.
4. Model and rehearse the appropriate actions, emotional responses and self-control.
5. Teach clients that dysfunctional responses are cues for the implementation of coping skills.
6. Assign ways to practice imaginal and in-vivo responses to stressful situations that become increasingly difficult over the course of therapy, so as to foster confidence and the learning of skills.
7. Promote greater insight into one's coping abilities and potential.

The three phases of stress inoculation training are *conceptualisation*, *skills acquisition and rehearsal*, plus *application and follow through* (Meichenbaum, 1985). The conceptualisation phase is primarily centred around establishing a sound therapeutic relationship, educating the client about the transactional philosophy of SIT, and defining as precisely as possible the target behaviours and the ongoing stressful events of concern. The skills acquisition and rehearsal stage involve the actual coping skills training components, and the gradual elimination of inhibitory responses and fears. The application and follow-through involves the applied use of the skills learned, the fostering of an independent attitude and the recognition that SIT skills need to be maintained in an ongoing fashion. The application and follow-through phase also typically involves booster sessions and rigorous follow up procedures.

Closely related to SIT is problem-solving training (PST). Although PST can potentially be seen as an independently developed cognitive strategy (see D'Zurilla & Goldfried, 1971), I have included it under the umbrella of Meichenbaum's approach because most SIT related approaches use PST as a major component strategy (Meichenbaum, 1977, 1985). PST is aimed at encouraging the client to deconstruct a problem into a manageable sequence of behaviours when faced with a difficult situation (Meichenbaum, 1985). Client's are first shown how to identify the problem (as a stage requiring attention in itself), and then create a list of efficient ways to handle the problem and predict the outcomes for each alternative. After generating these alternatives, a cost-benefit analysis dictates which ones will be selected and tried. Meichenbaum (1977) offers a synthesis of the different types of problem-solving training in terms of teaching a client to follow a sequence of nine steps:

1) Clearly define the problem.
2) Create a list of realistically attainable outcomes.
3) For each possible outcome, detail the corresponding course of action that it requires.
4) Compare these plans with some solutions you might expect other people to try in the same situation, and add these to your own list of alternatives if necessary.
5) List the benefits and disadvantages of each alternative, and proceed to rank them in order of preference on this basis.
6) Rehearse the alternatives through imagery or behavioural practice.
7) Try out the alternatives that seem most beneficial and most likely to achieve the desired outcome.
8) Expect failures, but reward oneself for trying. Repeat the procedure if the initial solution was found to be ineffective.
9) Remember the original situation as amenable to the problem-solving strategy for future reference.

In addition to PST, SIT also uses exposure, relaxation training, cognitive restructuring methods similar to REBT and CT, as well as homework assignments and other operant conditioning procedures. In general, SIT appears to be well suited to relatively specific

problems. To date, empirical studies testing strategies which employ Meichenbaum's (1977) model of SIT have found the approach be more effective than wait-list control and as effective as assertion training and supportive therapy in the treatment of post-traumatic stress disorder (PTSD)(Resick, Jordan, Girelli, Hutter, & Marhoefer-Dvorak, 1988); in combination with PST, SIT is more effective than relationship therapy and attention placebo in treating conduct disorder (Kazdin, Bass, Siegel, & Thomas, 1989); more effective than a control group in treating impulsivity problems in children (Meichenbaum & Goodman, 1971); and superior to control in chronic pain management, including arthritis (Keefe et al., 1990), abdominal pain (Sanders et al., 1989), and back pain (Linton, Bradley, Jensen, Spangfort, & Sundell, 1989). Earlier research also showed preliminary support for the effectiveness of SIT in managing anger (Novaco, 1979) and the reduction of test anxiety (Hussian & Lawrence, 1978).

2.5 The Social-Learning and Social-Cognitive Theories of Albert Bandura.

The work of Albert Bandura has received much attention over the last two decades, most notably for his theory of *self-efficacy* (Bandura, 1977a) and his related and broader theory of *social-learning* (Bandura, 1977b). Bandura later integrated these two areas and his approach is now most commonly referred to as *social-cognitive* theory (Bandura, 1986). I have left Bandura's approach until last because it is somewhat different from the preceding three cognitive schools. Unlike, Beck, Ellis, and Meichenbaum, Bandura's approach is not so much a bona-fide therapy per se, but rather an attempt to create a general theory of human functioning that is amenable to both cognitive and behavioural terms. Bergin and Garfield (1994) identify four general points as being central to Bandura's (1986) social-cognitive theory:

1. Change occurs in a social context and is primarily cognitive.
2. Behaviour is generally self-regulated by the person as an agent within a context of reciprocal causality.

3. The estimations that people make of their own abilities is one of the best predictors of their actual behaviour.
4. Behaviour is principally controlled by forethought, reflection, interpreted meaning, and locus of control.

Bandura (1986) argues that the environment is not independent of behaviour, because environmental events themselves must be dependent on the situations people seek out. In turn, because our actions alter the environment, cognition can not be seen as the sole determinant of behaviour. Thus, a central idea to Bandura's approach is an emphasis on *reciprocal determinism*, the fact that behaviour itself can alter the variables of which it is a function (i.e., aspects of the environment). Bandura also claims that learning is not always directly experiential; we often learn new behaviours and ways of thinking by observing the people around us. Thus, another central idea behind Bandura's (1977b) approach includes the phenomenon of *social modelling*, the innate tendency for people to imitate others.

Bandura (1986) holds that the interplay between environment and human action is mediated by what he has called the *self-system*. The self-system consists of cognitive schemas which organise our self-evaluation, self-regulation, and self-perception. The self-system exerts control over behaviour by creating *self-efficacy* beliefs. A self-efficacy belief has been defined by Bandura (1977a) as "the conviction that one can successfully execute the behaviour required to produce the outcomes" (p. 193). From this viewpoint, behaviour is not always best predicted by a person's intrinsic capabilities, or even past behaviours, but by one's *perceived expected performance*. Of course, self-efficacy beliefs are necessarily influenced by one's abilities and experiences. However, a person's capacity to learn *new* behaviours is itself influenced by her self-efficacy. Thus, self-efficacy is reciprocally-determined in that beliefs can impact on behaviour which can in turn affect the environment. Bandura (1977a; 1986) has identified four major sources from which self-efficacy beliefs arise:

1. ***Mastery Experience/Performance Accomplishments:*** These are self-efficacy beliefs which are constructed on the basis of one's own past performance. They have,

according to Bandura (1977a), the greatest influence on self-efficacy expectations. However, Bandura emphasises that efficacy attributed to one's achievement will vary enormously across different people, and thus performance itself is not a sole determinant of these beliefs. However, a useful way of inducing behaviour change might involve first starting with non-demanding tasks as target behaviours, specifically designed to evoke a positive level of perceived competence (e.g., ensuring success by using graded homework assignments that start at a very basic level).

2. ***Verbal Persuasion:*** Ensuring that success for the client is attainable is not always sufficient on its own. A second source of self-efficacy information is based on the direct verbal feedback given by others. This can be an important source of behavioural change, particularly in therapies which are primarily semantic (e.g., psychoanalysis and Rogerian approaches). Any verbal interaction that results in an increase in appropriate self-efficacy can be considered valuable, and strategies used to eliminate sources of persuasion which defeat or weaken positive self-efficacy beliefs should be undertaken. Bandura (1977a) states that this last point is important because it is much easier to diminish self-efficacy through poor appraisals than to increase it with positive appraisal.

3. ***Physiological Arousal:*** People use their states of emotional arousal, affective conditions, mood, and other physical symptoms to evaluate their performance level. For example, a state of relaxation before a public speech is a good indicator that one will do well, and may be used to construct a belief regarding one's public presentation. Furthermore, in keeping with Bandura's emphasis on reciprocality, social-cognitive theory holds that self-efficacy beliefs can also contribute to physiological states, completing the belief-environment loop. For example, if a person were to conclude on the basis of their heart rate that they were likely to perform poorly when giving a speech, this could worsen anxiety levels (e.g., an increased heart rate). People have a natural tendency to monitor their physiological

states and from this information create beliefs about how effectively or poorly they will perform.

4. ***Vicarious Experience:*** Vicarious experience refers to self-efficacy beliefs which are based on our observation of others. In line with the role of modelling in social learning, Bandura holds that a person's evaluation of other people's performance, particularly in relation to their own, is an important source of self-efficacy beliefs. Although generally less influential than direct experience, in certain conditions vicarious experience can have the most powerful influence on one's self-perceived performance. We are particularly likely to use the performance of another as a basis for comparison if that person is regarded as highly competent. For example, a student who judges her writing skills in relation to her teacher would probably have a more stringent self-evaluation than a student who compares her writing to fellow students.

According to Bandura (1977a), the self-system is a major determinant of personality. Self-efficacy beliefs influence the choices we make and determine the tasks in which we will engage. For example, we are more likely to avoid tasks for which we have low self-efficacy. Furthermore, self-efficacy beliefs range from being quite general to those which are highly task-specific. While one person might believe themselves to be poor at preparing food in general, another may believe that "baking sponges is the only thing I'm hopeless at". Furthermore, in everything we attempt, self-efficacy beliefs are likely to determine how long we will persevere with a task, how much adversity we will endure before a task is abandoned, and how much effort we will allocate to a task. Thus, self-efficacy beliefs probably play a large role in motivation. Perhaps the role self-efficacy beliefs play in one's personality are most salient when one considers that the interpretation of task difficulty is likely to have a profound influence on one's opinion of the world; if someone consistently thinks poorly of their ability, it follows logically that they will be more likely to perceive the world in a negative light. Thus, two students can be engaging in the same tasks and judged by observers to be achieving equal success. If levels of self-efficacy were discrepant between the two students, one might assess the task as "an achievement to be had", while the other may view it as "a failure to avoid".

Although sometimes difficult to separate, Bandura (1986) believes that predictions of performance differ from "judgements of the likely consequence that behaviour will produce" (p. 391). Thus, when trying to identifying the nature of a client's self-efficacy beliefs, therapists must distinguish between interpretations of the likelihood of successful behaviour, and interpretations of the unavoidable consequences that behaviour might produce. Of course, the two constructs are intertwined in that the likelihood of reward-successful behaviour is highly dependent on the prevailing contingency.

As mentioned before, Bandura's social cognitive-theories do not seem to be typically used as a mechanism of psychotherapy per se, although there are direct applications of self-efficacy research, most notably in academic settings (see Pajares, 1996). However, practitioners of CBT may use a range of techniques which operate on Bandura's constructs by aiming to increase self-efficacy and attending to the social context and reciprocal effects of dysfunctional behaviour. As such, self-efficacy beliefs have been found to be a critical factor in the development and amelioration of phobias (Bandura, 1983), depression (Davis & Yates, 1982) social skills and interpersonal effectiveness (Moe & Zeiss, 1982), addictive behaviours (Marlatt, Baer, & Quigley, 1995), the management of chronic pain (Manning & Wright, 1983), and various other psychological maladies (for an analysis of the use of social-learning theory within CBT programs, see Zarb, 1992).

2.6 Interpreting Cognitive Therapy as RGB: The Field of Psychotherapy Integration.

Any effort to define cognitive therapy in radical behavioural terms can be seen as a task belonging to the field of *psychotherapy integration.* As a field of inquiry, psychotherapy integration looks beyond the epistemological boundaries of various psychological schools in order to gain new knowledge applicable to methods of behavioural change. There are a number of reasons why researchers have called for a review of the boundaries dividing the many modes of therapy that now exist, estimated to total as many as 400 (Karasu, 1986). Among these reasons include the inefficiency of single school adherence, the equal outcomes across therapies phenomenon, and the possible advantages of having a universal language in clinical psychology (Goldfried,

1983). Integrative work may be directed towards creating more comprehensive and accurate theories of psychopathology, or directly altering and improving the formulation and administration of psychological treatment programs. The integrative movement has been paired with the gradual appearance of eclectic thinking among psychotherapists, and paralleled a growing negativity towards the rigidity of single school approaches.

The first attempt of an integrative analysis was probably the study initiated by French (1933) which addressed the overlap between psychoanalysis and theories of respondent conditioning. Like French, the majority of integrative work which took place in the subsequent three to four decades was also chiefly concerned with psychoanalytic and behavioural rapprochement, the most influential of which was undoubtedly the seminal attempt by Dollard and Miller (1950). During these times, the field of psychotherapy integration was largely driven by a small collective group of researchers who queried the traditional restrictions placed on current fields. Although this was a relatively quiet movement, it laid the groundwork for the strong and coherent force that psychotherapy integration has become over the last two decades (Arkowitz, 1992). This surge of interest is reflected by events such as the establishment of the *Society for the Exploration of Psychotherapy Integration* (SEPI, est. 1983), and the appearance of two journals devoted to integration, the *Journal of Eclectic and Integrative Psychology* (est. 1992) and SEPI's *Journal of Psychotherapy Integration* (est.1991). Additionally, there is now an increasing number of books entirely focused on the merits of integrative thinking (e.g., Clarkson, 1997; Feltham, 1997; Gold, 1996). Most notably, however, the vast growth in the overall number of published studies which include eclectic and cross-paradigmatic thinking can be taken as evidence of major attitudinal changes within psychotherapy research.

According to Arkowitz (1989), psychotherapy integration can be roughly divided into three major fields: *technical eclecticism, common factorism,* and *theoretical integration. Technical eclecticism* is the least theoretically based of these three, and was first introduced by Lazarus (1967). Although technical eclectics work within one predominant framework they freely borrow from other orientations when circumstances dictate, providing any such techniques have obtained empirical support. Probably the most prominent example of the technically eclectic approach is *multi-modal therapy*, designed by Arnold Lazarus himself (Lazarus, 1981). Multi-modal therapy is often recognised by

its popular acronym; "BASIC ID", a tool used to guide clinical assessment according to seven major areas of psychological makeup (i.e., behaviour, affect, sensation, imagery, cognition, interpersonal relations, & drugs/biology).

Common factors integration refers to the identification of elements shared between two or more modes of therapy. It is hoped that by obtaining clarification of the common effective ingredients that contribute to a positive outcome, researchers may be in a position to isolate the mutual ground shared by fundamentally different therapies. A classic example is Frank's (1961) book "Persuasion and Healing". Frank (1961) argued that the healing effects of almost all therapies are the direct product of a clinician's ability to influence a range of inducible states such as hope, cathartic releases, self-understanding and an enhanced sense of mastery. An example of a more contemporary common factors model can be found in Karasu (1986) in which all therapies are said to provide varying elements of cognitive mastery, affective experiencing, and behavioural regulation. Karasu establishes the importance of these three elements by identifying how each of the most popular therapies function around these constructs.

Theoretical integration refers to rapprochement via epistemological and ontological compatibility, or the synthesis of frameworks in which two or more paradigms can co-exist (Arkowitz, 1989). Proponents of theoretical integration aim to re-evaluate the most essential aspects of the philosophical stance held by psychotherapeutic modes. It is hoped that a closer evaluation of these core boundaries will reveal superficialities in the differences between two or more paradigms. At this level, theoretical integration is probably directed towards developing new research strategies as much as new forms of clinical practice.

Schact (1984) has also defined several other forms of psychotherapy integration. *Translation* refers to purely redefining the techniques of one modality in the terminology of another (i.e., similar to common factors integration, only the common factors are defined by the nomenclature of *one* of the two integrated schools). *Synergism* refers to the application of two techniques to the same problem in the hope that there will be greater overall client change than that possible with a single application of either technique. *Complementation* uses a range of theoretically different and specific techniques, each identified as most appropriate for a particular problem and combined

accordingly into a structured treatment package (i.e., similar to technical eclecticism, only the appropriateness of the problem-technique match determines inclusion, rather than empirical support). The *emergent model* involves the blending of two philosophies to produce a unique and new therapeutic approach, containing almost none of the singular characteristics of its constituent therapies. Finally, Schact (1984) refers to *theoretical* integration as the process of reviewing and modifying the epistemological boundaries of two therapies to produce a converted single approach, as in the *theoretical integration* method identified by Arkowitz (1989) discussed earlier. Schact (1984) also points out that while some integrationists focus entirely on therapeutic techniques, others focus on theories of psychopathology, client change, or even the unification of different models of personality. Thus, as there are many types of rapprochement, and an enormous number of psychotherapies that could potentially be combined, there is a myriad of potential integrative projects available to the researcher.

It is hard to identify which form of psychotherapy integration could accurately represent an immersion of cognitive therapy into behavioural terms using the concept of RGB. It seems that the distinction made by Arkowitz (1989) between technical eclecticism, common factorism, and theoretical integration reveals itself to be somewhat arbitrary upon closer examination. Separating common factors from theoretical integration seems at times impossible. For example, one could argue that the goal of increasing social activity is a factor common to both the cognitive and behavioural schools in their treatments of depression. But "increased social activity" belongs to the school of behaviourism if it has come about through reinforcement, but belongs to the cognitive school if it is a product of altered beliefs about one's social value. Thus, it is sometimes unclear at which level common factors are to be conceptualised before they are awarded or denied the status of "theoretical". A simple way to resolve this conflict is to contend that common factors are never completely atheoretical and simply can not be detached from their philosophical roots when identified as similar to the process of change in another therapy. Finally, the difference between eclecticism and common factorism can not be clearly defined in some cases. Common factors are always identified

in pursuit of a possible enmeshment of practical techniques,[8] the final combination of which can almost always be construed as some type of eclectic practice (technical or not).

Schact's (1984) categories seem slightly more promising. The concept of integrative *translation* seems to reflect some of the elements of the type of integration sought here. However, a mere translation would imply that there are no hybrid components derived from such rapprochement. As mentioned in Part One, discovering promising *new* variables within a behavioural analysis is a major goal of a cognitive-behavioural integration based on the concept of RGB. To overlook this would render the current project as simply wanting to semantically paraphrase cognitive techniques using operant terms, for no apparent reason other than to satisfy the epistemological whims of ardent behaviourists. Another goal of a rule-governed interpretation of cognition is to completely avoid the need for information-processing metaphors and logical analogies; this would therefore also exclude *synergism*, where different therapies *retain* their philosophical roots but are applied to the same client problem in the hope that additional effects are seen. *Complementation* does not seem to fit either; this would imply that the new integrated model should retain cognitive concepts where superior effectiveness is found, and this would not be amenable to a *pure* radical behavioural interpretation. The *emergent* model is inapplicable also, as an RGB integration is designed entirely in *order* to retain the core concepts of radical behaviourism (i.e., rather than completely removing virtually all of the fundamental tenets of the integrated therapies, which is typical of the emergent approach). Thus, none of the aforementioned categories of integration are viewed here as an adequately defined guiding theme for an RGB cognitive–behavioural integration.

One conceptualisation of psychotherapeutic rapprochement which could be helpful here is Messer's notion of *evolutionary* or *assimilative* integration (Lazarus & Messer, 1991). The concept of assimilative integration is based on the observation that paradigms tend to integrate naturally as they expand and eventually come to incorporate previously foreign concepts and phenomena (Messer, 1986). Messer proposes that paradigms in

[8] Setting aside, for the moment, researchers who identify common factors for the purpose of stimulating new hypotheses and intellectual debate.

therapy can and do integrate through a natural and selective importation of one therapy's concepts into the framework of another (Lazarus and Messer, 1991):

> I am incorporating an attitude, perspective, or approach that is transformed in its new context even while retaining something of value from its point of origin. It then requires validation through clinical use and experimentation, just as might any new therapeutic modality (p. 153).

Messer (1986) argues that the selective importation of concepts must be carried out in such a way that they fit comfortably into the larger theoretical context, whether initiated by a recognition of implicit similarities or otherwise. Messer claims that instead of importing techniques and observations shorn of their theoretical contexts, psychotherapy integration should occur by importing a concept or approach that retains something of its former theoretical value in the dominant framework. If we take radical behaviourism to be the predominant framework here, then we can see the importation of cognitive strategies as being possible through the "naturally expanding diversity of radical behaviourism". By adopting the assimilative philosophy, importance is placed on whether an RGB C-B integration retains the value of the cognitive strategies it reframes, while detracting attention away from theorists who would perhaps prefer to see RGB couched in cognitive terminology, or the construction of an amalgamation of concepts from both schools. Furthermore, an assimilative integration can potentially meet the goals of each of common factorism, theoretical integration, and technical eclecticism; the assimilative approach aims to find those factors that are similar by expanding the theoretical boundaries of behaviourism, especially those which previously prevented the immersion of a "cognitive" phenomena, in the hope that the formulation of a new treatment strategy can be produced.

To summarise this subsection, it is argued that the assimilative approach is a *valuable philosophy to theorists and therapists wishing to rationally extend the practical applicability of their school.* Because of the logical deficiencies and obscurities inherent in other integrative methods, and because the assimilative approach encourages the researcher to (a) maintain her philosophical stance, (b) seek empirical support for the

integrated model, and (c) search for new hypotheses and variables, it will be adopted here. Accordingly, from this point onwards I will refer to an RGB C-B integration as *assimilative*. We now turn to Part Three, where the assimilative approach will be applied to the cognitive and behavioural therapies using the concept of RGB.

PART THREE

RULE-GOVERNED BEHAVIOUR AS AN APPROACH TO COGNITIVE THERAPY

In the third and final section of this thesis, I discuss how several core components of cognitive therapy might be evaluated in terms of the concept of RGB. Because the emphasis is on cognition, the analysis will be primarily focused on self-rules. However, I will first outline Zettle and Hayes' (1982) operant analysis of the *verbal event*, as there are several ideas introduced under this heading that will be relied on throughout Part Three. I will then identify three forms of dysfunctional rule-following. There are probably others, but I will only discuss what I have termed *intraverbally designed self-tracks*, *inadvertently maintained rule following*, and *overly rigid pliance*. These three are the most relevant to the core terms and concepts outlined in the introduction. I will then discuss several dysfunctional rules that can arise as a result of faulty rule-construction, and the possibility of creating disorder-specific cognitive theories of psychopathology within the framework of RGB. Each of these areas will be addressed in light of their relationship to the different cognitive therapies outlined in Part Two, and to a lesser extent, cognitive therapy in general. Following this, I shall discuss cognitive theories of personality, and introduce a flow diagram which might be used to assimilate some of the qualities of Bandura's (1986) self-system. I will then outline the major mechanisms of change that a rule-based therapy could employ, and how these relate to cognitive mechanisms of change. This will be followed by a brief discussion of how cognitive techniques interpreted as RGB could be effectively combined with traditional behaviour therapy. I will then discuss some of the future empirical and theoretical problems confronting the concept of RGB, and their relationship to the present essay. The final subsection includes some generalities about the field of psychotherapy integration and its relation to the current project. My concluding comments are then submitted.

In order to recap on the introduction, the major verbal units presented so far have been tabulated below (Table 3.0). These will be the major units of analysis in Part Three.

Table 3.0 Fundamental units of speaker and listener verbal behaviour.

Verbal Unit	*Definition*
Verbal Behaviour	Behaviour reinforced through the mediation of others.
Tact	Speaker behaviour under the control of an environmental stimulus.
Mand	Verbal behaviour of the speaker controlled by a state of deprivation or aversive stimulation.
Autoclitic	Speaker behaviour which operates as a function of its ability to alter the function of an accompanied verbalisation.
Intraverbal	Speaker behaviour which is in response to the presence of other verbal behaviour.
Precurrent Behaviour	Behaviour which generates, and is reinforced by, the presence of a discriminative stimulus which may serve to occasion reinforceable behaviour.
Pliance	Listener behaviour operating as a function of the apparent correspondence between the rule and the contingencies mediated by the generator of the rule (i.e., the ply)
Tracking	Listener behaviour operating as a function of the apparent correspondence between the rule and the environmental contingencies specified in the rule (i.e., the track)

3.1 The Verbal Event: Meaning and Errors of Translation.

Central to cognitive therapy is its analysis of *meaning*. Beck (1976) has emphasised the importance of the way people attribute meaning to events, and has expressed his dismay at the way behaviour analysts tend to "avoid the issue altogether" (p. 148). The word *meaning* can be defined as "that which is intended" (Chaplin, 1985). Taken literally, the word *intended* normally implies that something other than the face value of an event is to be conveyed to the designated receiver of such meaning. Perhaps then, the

issue of meaning is at its most complex when trying to understand human interaction. The way people "interpret" what they say to themselves and what other people say to them undoubtedly plays a major role in human functioning. To the radical behaviourist however, the process of expressing and understanding meaning can be understood more scientifically in terms of verbal behaviour; the fundamental unit of which is the *verbal event*. Zettle and Hayes (1982) have offered a detailed analysis of the verbal event and RGB that may assist in an explaining the radical behavioural fabric of cognitive therapies, and how we might conceptualise "that which is intended" from a behavioural perspective.

Zettle and Hayes (1982) analysis of the verbal event emphasises the distinction between formal and functional definitions. As explained earlier, any event which can be described functionally can always be described formally also. By *formal*, I mean a unit that has been described according to its structural qualities, rather than its function. For example, the behaviour of answering a knock at one's door might be functionally described as "behaviours which serve to open the door and are occasioned by a knock". However, we may also describe the exact structural features of any instance of this behaviour, such as "using one's hand to turn the doorknob", "walking towards the door", "turning the dead-lock lever clockwise", and so forth. This would be a description of the formal qualities of this event. It was mentioned earlier that tacts often appear as descriptions while mands often look like requests. This too is a reference to structural properties.

We might also try and formally describe what a track or ply *typically* looks like to the verbal community. Because plies function by implying speaker-mediated consequences for rule following, they often contain formal units such as "can you please...", "I want you to...", or "go and do X, or else I will Y...", and other words which we might casually describe as "conveying an appeal" (i.e., mand-like units). Alternatively, because tracks function by specifying an environmental contingency of interest to the listener, they sometimes contain more descriptive-like formal units, such as "this is a..." , and "anyone who does this will receive..." , or "If you do X, Y will occur...", and so on (i.e., tact-like units). It is also important to understand that a rule can be made up of several different tacts or mands, and as a compound, may first occur in completely novel

combinations. As Catania (1998) has pointed out; "even if you've never seen a purple cow, the separate tacts of *purple* and *cowness* will allow you to say "that is a purple cow" when you encounter one" (p. 257). Thus, both tacts (and sets of tacts) and mands (and sets of mands) tend to function as tracks and plies, respectively. The statement "that is a purple cow" might function as a track to a nearby listener – even a listener with no prior experience with such a contingency would probably look around with the expectance of being reinforced with seeing a purple cow (providing the rule came under the class of those instructions which we might describe as specifying "reinforcement for looking at something"). However, these formal components, as typical as they may be, are only incidental in comparison to the functional properties of a rule.

Accordingly, Zettle and Hayes (1982) have argued that the verbal antecedents which occasion RGB can be deceptive because of the way their formal and functional properties can alternate. As both tacts and mands are defined by the speaker's individual history of reinforcement, identifying the two at any one point can be difficult without knowledge of the contingencies surrounding the speaker. Although a tact will often *appear* to the listener in a descriptive form, typical of the tact, it may function purely as a mand. For example, upon hearing a telephone ring someone might say "the telephone is ringing" – such a statement is in tact form, in that it simply appears to be controlled by the presence of the telephone and a history of reinforcement for saying "the telephone is ringing" in the presence of a ringing telephone. It simply appears to describe a natural event. However, such a statement may *function* more like a request. This person may have said "the telephone is ringing" with the intention of using such a statement to stimulate a nearby listener to answer it.

Functionally defined then, this verbalisation may have met the criteria for a mand, because it was stated in order to control the behaviour of the listener[9]. Tacts can also mimic the typical formal features of a mand. For example, in the presence of a ringing telephone, someone might say "if nobody answers that I'm going to be really angry". This appears to be a mand, because it appears as though it is commanding a nearby listener to answer the telephone. That is, at least some members of the verbal community

[9] Again, however, it is entirely possible that saying "the telephone is ringing" could also meet the functional criteria for tacting if the appropriate conditions were met.

would probably answer the telephone in the presence of such a statement. However, this could be a tact from a functional perspective, in that the speaker has not actually said it so as to stimulate a listener to answer the telephone, but is simply tacting the current state of affairs. In this case, the speaker is tacting both her volatile emotional state and the ringing telephone. Thus, because both tacts and mands can occur in forms characteristic of each other, the *structural forms of either may be indistinguishable, but not with regards to their controlling variables* (i.e., not from a functional perspective).

Thus, any verbal event is characterised by a *functional* unit of speaker behaviour, a *formally* described speaker unit, and a *functional* unit of listener behaviour. On most occasions, tracking will be in response to a tact which appears as a tact, and pliance will most commonly be in response to a mand which appears as a mand (Table 3.1). However, it is not the case that everyone simply and explicitly says what they want, nor is it the case that everyone does what they and others have told them to do. Speakers may manipulate the form of speech for a multitude of reasons, and whether a tact or a mand will produce pliance or tracking will depend entirely on the listener's experience with such utterances. For example, a mand may appear as a tact and produce tracking (Table 3.1). Imagine a parent wanting to control a child's behaviour with the rule "be nice to your brother or Santa will not come!". This would be a mand if it was controlled by the effects it was to have on the listener, which is the child. The parent has not tacted components of a contingency about arguing and receiving presents from Santa because such a contingency does not exist, and like all mands, if such a statement usually had no effect on the listener (child) the speaker (parent) would not have said it. However, *to the listener* this statement would probably *appear* to be in typical tact form because it gives the impression that the parent has at some point tacted a situation in which Santa's contingencies were in effect. That is, to the child, this statement appears to describe the "fact" that receiving presents on Christmas day is contingent on ceasing to argue. As such, the child probably responded to (or tracked) the rule as though the description had an accurate correspondence with the natural contingency pointed to (track). Thus, a speaker's utterance can be a mand *masquerading* as a tact, and because of its recognisable form as a tact, would probably evoke tracking. Zettle and Hayes (1982) call these units *trojan tacts,* because like the trojan horse, their form conceals their true

nature. Imagine if this tact had produced pliance anyway. Perhaps what the child actually "heard" was her parents *intended* function of the rule, as though it was more like "stop arguing with your brother, or else!". In this case the tact would not be a trojan tact, because it would have failed to hide its true function as a mand. Trojan tacts are probably most commonly used when the speaker desires to alter the listener's behaviour, but needs to hide some ulterior motive.

In the same manner, a tact may still appear as a tact, but produce pliance (Table 3.1). For example, a customer of some restaurant might say to the chef "this soup has a lot of paprika". Imagine that this is a tact and its occurrence is no way intended to stimulate an alteration of the behaviour of the listener (the chef). Indeed, the customer could simply be making a passing comment about the flavour of the soup, in the presence of the chef. However, the chef might respond to such a statement as a threat, as if the statement was more like "if you don't give me another bowl of soup with less paprika, I will leave this restaurant". Now imagine this chef had seen customers leave on other occasions after making such comments, but had also seen unsatisfied customers *stay* when their food was replaced. If this was the case, then this tact might function as a ply because the chef would have responded as though the statement indicated the possibility of aversive speaker-mediated consequences for non-compliance (e.g., the speaker leaving without paying). To comply with this rule then, the chef might offer a bowl of soup with less paprika, in order to obtain the negative reinforcement of preventing the patron leaving. This would be the rule governance of pliance, because it developed out of following customer-supplied rules demanding "good restaurant service". Zettle and Hayes (1982) have named this reactive-like sequence *counterpliance* (Table 3.1).

Many other combinations, not discussed by Zettle and Hayes (1982), also seem possible. For example, a tact could masquerade as a mand and produce pliance, what we might perhaps call a *trojan mand* (Table 3.1). A trojan mand might simply be a general statement of fact (i.e., originating through tacting), but inadvertently imply a demand or request of the listener because of its mand-like form. Furthermore, a mand could appear as a mand but produce tracking, what we might perhaps call *countertracking* (Table 3.1). Countertracking would take the form of requesting something in a request-like form, but "anticipating" that the listener will follow it because it specifies some other contingency.

However, mentalistic terms such as *anticipate* are unnecessary when one accounts for the extraneous contingencies that might produce tracking.

Table 3.1 Sequences of verbal episodes.

Functional Speaker Unit	Structural Appearance (Form)	Listener Behaviour (RGB)	Sequence Name.
Mand	Mand	Pliance	Normal RGB
Tact	Tact	Tracking	Normal RGB
Mand	Mand	Tracking	Countertracking
Tact	Tact	Pliance	Counterpliance
Mand	Tact	Tracking	Trojan Tacts
Tact	Mand	Pliance	Trojan Mand

Zettle and Hayes (1982) have shown that an analysis of the contingencies operating on both the speaker *and* the listener are essential in order to visualise how the two can be at cross-purposes in any one verbal event. This is perhaps what cognitivists would refer to as the thorny issue of defining misinterpreted meaning. To the RGB C-B therapist, analysing verbal-episodes in this way might be most useful in the assessment of problems of social interaction. However, the way people interact with themselves will also involve these sequences, and this appears to be a central focus of cognitive analysis. As will be demonstrated, these structural-functional distinctions and interactions can be usefully employed in a behavioural analysis of disorders of rule-following and formulation.

3.2 Disorders of Rule Following I: Intraverbally Designed Self-Tracks

I will now introduce the first of three disorders of rule-following and their relation to cognitive therapy. As mentioned earlier, incorrect tacting is usually rapidly punished by one's verbal community, especially during situations where a tact is expected by others. If I were to say "that's a lovely hat" when someone had given me a tie, that person might say "What?, that's a tie!". Similarly, tracking is usually quickly extinguished if the track

is found to be incorrect and subsequently identified as faulty. If I were inexperienced enough to follow a clearly specified track such as "eat your eggs with dirt and you will experience a great taste sensation", I would be punished with the taste of dirt for following it, and would probably never follow that track again. I would not follow it again because next time I was given that rule, by myself or someone else, I would respond to the rule as resembling one with little or no correspondence with a reinforcing environmental contingency (i.e., I would identify it as not involving sufficient tacts).

Problems can occur, however, when self-tracks are presented in a vague or *untestable* form. Recall, to self-track is to follow a self-rule because it appears to correctly specify an environmental contingency. Zettle and Hayes (1982) have suggested that self-tracks may be untestable when they employ the use of *intraverbals*. An *intraverbal* is a verbal behaviour in response to other verbal behaviour, as opposed to having been directly associated with non-verbal environmental stimuli or objects (Skinner, 1957). Thus, if a track involves an intraverbal, it is often unlikely that it can be tested for its rule-environment correspondence because it is probably too vaguely connected to environmental stimuli. For example, take the man who says to himself "it is terrible when other men look at my wife". Although the words *wife* and *men* are tacts of environmental stimuli, the word *terrible* is not. It is not a tact because the verbal community was not likely to have repeatedly reinforced this speaker for saying the word *terrible* in the presence of other men looking at his wife. Instead, the word *terrible* is an intraverbal, probably privately derived, and only thematically associated with this speaker's wife and other aspects of the troublesome situation.

Nevertheless, as Zettle and Hayes (1982) point out, because of their descriptive or tact-like form, tracks which involve intraverbals *may still cause the rule-governance of tracking*. As a track, the statement "it is terrible when other men look at my wife" may evoke RGB such as avoiding places where other men might be. Had the rule been in a more testable form, specific and devoid of intraverbals, then the ensuing dysfunctional rule-control might extinguish. For example, take a more adaptive rule such as "even if I take my wife to a place where other men can look at her she would still tell me that she loves me if I were to ask her". This rule could be easily tested for its rule-contingency correspondence because the rule would easily be discriminated as correct (or incorrect)

upon following it. As I will explain later, this husband might have benefited from formulating the more adaptive rule by seeking out such a contingency and describing it. It should be noted that the very use of intraverbals themselves is unlikely to be problematic. For example, the rule "I am an amazing netball player" is likely to produce tracking which is very adaptive, even though the word *amazing* is an intraverbal (e.g., increased confidence on the netball court). It is only when such statements cause undesirable RGB that they are problematic[10]. Here I refer to such tracks as *intraverbally designed self-tracks*.

Much of Ellis' (1962) REBT may centre around uncovering intraverbally designed self-tracks (Poppen, 1989; Zettle & Hayes, 1982). For example, Ellis (1995a) speaks of "awfulising", and the tendency for clients to "catastrophise" events. The word *awful* is intraverbally derived also, having not been learned through a direct association with any stimulus event. Thus, the statement "It is absolutely awful to be embarrassed" is untestable, yet because of its descriptive tact-like form it may still produce tracking; with this example, perhaps the avoidance of potentially embarrassing situations. Another irrational belief might be the statement "it is pitiful to cry". The word *pitiful* is also an intraverbal and is not experientially associated with any specific contextual stimuli, as tacts by definition must be. Nevertheless, a "repressive" behaviour, such as refraining from crying, may be the ensuing tracking. Ellis would probably label such self-restraint as being "irrational".

Beck also focuses on the problematic use of such intraverbals. A particularly common one is the adjective *worthless*. Beck (1976) has argued that a prominent feature of depression is a feeling of 'worthlessness'. Presumably, such a belief is at some point expressed or verbalised as a statement such as "I am worthless" (if not, how could Beck have come to this conclusion?). From a radical behavioural perspective, it is easy to see that the statement "I am worthless" could have drastic effects on one's life if it were to function as a rule. It is altogether likely that such a rule would function as though it implicitly specified contingencies such as "If you try and love somebody they will not appreciate you", or "the only way to escape the existence of a worthless person is to kill

[10] This emphasises my point that the three disorders of rule following discussed here are unrelated to my later analysis of dysfunctional rule-*construction*, where it is clearly the structure of the *rule* that is at fault.

yourself". Unfortunately, such a rule is likely to persist as a dominant component of the depressed individual's verbal repertoire, because the intraverbal *worthless* prevents the rule from being testable - what possible behaviour could someone engage in that could wholly confirm or deny one's worthfulness? Or more definitely, what contingency could someone engage in that would illustrate a discrepancy between the track "I am worthless" and the actual state of affairs? As to be expected, often even significant achievements for the depressed person do not result in an abandonment of these detrimental tracks, perhaps what Bandura (1977a) might refer to as "negative appraisal", or what Beck prefers to refer to as evidence of the presence of "core schema". Thus, intraverbally designed tracks constitute a major part of dysfunctional cognition.

3.3 Disorders of Rule-Following II : Inadvertently Maintained Rule-Following.

In the introduction, I explained how rule-recital can be seen as a form of precurrent behaviour, the self-generation of stimuli that serve to evoke a reinforceable response. A rule is probably most often recited by oneself for no other reason than to evoke RGB. If I say to myself "try and be honest", and then do so, it is likely that I recited that rule purely as a way to exert some control over my own behaviour. However, the behaviour of rule-recital may be bound to contingencies *other* than those for rule-following. Self-rule recital may serve other behavioural functions in *addition* to the reinforcement of emitting a reinforceable response (i.e., in addition to being reinforced for serving the function of precurrent behaviour). Nevertheless, a rule may still be followed regardless of the multiple events or source of reinforcement that caused its presentation, just as any other verbal event can be causally related to multiple variables. Thus, problematic rule-following which occurs through additional contingencies for self-rule recital is labelled here as *inadvertently maintained.*

A simple way to appreciate the significance of inadvertently maintained rule-following is to use a non-verbal example as an analogy. Research on behavioural treatments for obesity have shown that certain situations exert powerful stimulus control over problematic eating behaviours (Weiss, 1977). One way to help reduce eating behaviours might be to reinforce the client for not entering high-risk eating areas at

inappropriate times. For example, not entering the kitchen and dining areas during periods other than designated mealtimes could eliminate one source of problematic stimulus control. However, there may be other contingencies which cause the client to enter the kitchen and dining areas, such as the need for social contact. This could be a crucial factor if the kitchen or dining areas were the only place where the client's family gathered. Thus, there may be contingencies of reinforcement for exposing oneself to (or for the self-generation of) certain stimuli that cause the repeated presentation of an unwanted stimulus. Similarly, it may be additional contingencies for the recital of dysfunctional rules that sustain their repeated presentation, and these may be relatively concealed in comparison to the behaviours actually *produced* by the rule.

What other contingencies might surround the behaviour of rule-recital? Skinner (1957) used the term *impure tacts* to refer to verbal tacts partly controlled by a state of deprivation or aversive stimulation in the speaker (i.e., partly controlled in the same way that mands are). For example, take the self-rule "always be honest about my occupation when meeting new people, as this will make me appear genuine and help me to make friends". Someone might understandably follow such a rule after accurately tacting an actual previous instance of meeting people (i.e., this person found on a prior occasion that such honesty ultimately gained people's trust and confidence, and subsequently described this relation). However, in addition to reinforcers for creating precurrent behaviour, imagine if this statement was also recited to relieve feelings of inadequacy about one's occupational status. If this were the case, then the contingencies for presenting an accurate track would be in competition with contingencies of anxiety relief. Thus, the tacts included in the track might be contaminated, or *impure,* but might still be used on later occasions as a rule. Zettle and Hayes (1982) claim that tracks which include impure tacts often still produce tracking because of their tact-like form. For example, the person who stated this rule might consistently make an effort to explain the merits of her job to people whom she has just met. Thus, tracks which are generated partly out of escape or avoidance may poorly tact or describe the contingencies to which they lead the listener to respond.

Meichenbaum often attends to this type of negative self-talk, as apparent in his following comment regarding SIT for rape victims suffering from PTSD:

> When the client responds with "I'm soiled property", or "I'm useless" or the words to that effect, it's usually said with a great deal of emotion. What I would do as a therapist is to not only attend to the client's words, but to the affect. So I would say "Soiled property, tell me about that." I would 'pluck' the key words and put them back in the client's lap, and encourage her to elaborate...After empathetically listening to their distress I might say something like: "Given what you've been through, if you were not depressed at times, if you did not feel at times like you were 'soiled property', or at times as if you were 'useless', then I think I would be really concerned, and conclude that something was seriously wrong." In this fashion, I attempt to validate the client's experience. I am going to go beyond that and even compliment the client for the symptoms of being depressed. For instance, I might say to the client, "What does your being depressed say about what's going on? ... Perhaps it conveys that you are in touch with your feelings, that you are reading your situation, that you're responsive to what you have experienced" (Hoyt, 1998, p. 54).

Alternatively, one can view Meichenbaum's attempt to "validate" the client's experience as a way of discovering the function that negative self-statements serve, over and above the ways that they may be directing behaviour. In the above quotation, it would appear that Meichenbaum is implying that statements such as "I'm soiled property" are reinforced with affective states describable as "being in touch with one's feelings", or feeling "responsive to one's experience". Presumably, Meichenbaum aims to elucidate these consequences to illustrate that they maybe inadvertently maintaining problematic behaviours. For instance, with Meichenbaum's example, this client might privately say "I'm soiled" when approaching a man for companionship, and this might evoke an abrupt cessation of this action. Because of the experiences and verbal repertoire of a rape victim, even a formal unit such as "I'm soiled" could function as though it were "I should not approach this man because he will eventually find out that I am dirty and I will get hurt".

Skinner (1957) specified that impure tacts can result from a state of deprivation in the speaker. The recital of many rules may be reinforced in this way. For example, Poppen (1989) points to Ellis' (1962) concept of "awfulising", claiming that clients may sometimes overtly state a rule such as "the world is awful" in order to get agreement from others. People within an audible distance of this client might realise that this is an appeal for empathy, and say things like "yes, there's so much wrong with the world today". What other members of the verbal community may *not* realise is that reinforcing such self-talk in the client may strengthen the use of rules which may later result in dysfunctional behaviour. For example, saying the self-rule "the world is an awful place" may function to evoke social withdrawal, or lowered motivation. Poppen uses the maxim "misery loves company" to describe this type of inadvertently maintained dysfunctional rule-following.

People may also inadvertently maintain their *own* rule-recital (Poppen, 1989). For example, a client says, "I know I'm going to lose if I play this game", and after losing say "I knew I was right", reinforcing oneself for constructing an accurate track. Thus, regardless of the behaviour that this rule might evoke, its recital serves to earn reinforcement from an additional source. In this case, self-reinforcement for the presentation of the rule was contingent upon its accuracy. Thus, we can see how self-defeating behaviour can be maintained by the reinforcing effects of confirmatory negative self-statements, and the hidden types of avoidance, escape, or positive reinforcement that they are contingent upon. Cognitive therapy may often be trying to remove problematic aspects of self-talk which is being *inadvertently maintained* by other factors.

3.4 Disorders of Rule-Following III: Overly Rigid Pliance

Tracks can produce schedule insensitivity if they in some way conflict with the natural contingency they point to, and this conflict is not noticeable (e.g., as with intraverbally designed tracks). For example, the Galizio (1979) study described earlier used participants with an FI avoidance schedule which randomly altered the position of

four different contingencies. The study found that rule-following which does not immediately result in punishment for inefficient responding will be maintained, even if it is not the most appropriate or sensitive method of obtaining reinforcement. However, this study probably involved a mixture of tracks *and* plies. There would have been times when the interval labels (e.g., 30 SEC) were followed because they appeared to correspond with the active schedule (tracking), but there were probably also times when they were followed simply because they were rules and there was no other evidence to go by (pliance). Indeed, it was most probably contingencies for pliance which produced the schedule insensitivity, because any discrepancy between the track and the natural contingency becomes apparent after tracking (unless the track is too vague or perhaps relies too much on intraverbals, which was not the case in the Galizio experiment).

Because pliance is primarily maintained by contingencies for *rule-following* per se, it is probably relatively insensitive to any natural contingencies to which it leads the listener to respond (Zettle & Hayes, 1982). That is, pliance is characterised by rule-governance which has been established through social contingencies for rule-following, and therefore reduces or limits the influence that other non-verbal contingencies can have on behaviour. Such insensitivity is probably useful at times. For example, I might say to myself when digging up a tree stump "I know I can do this if I try!". A ply such as this might help me to be insensitive to the punishing effects of extremely hard physical labour. Yet such a rule would only be followed because in the past I had been reinforced by speakers who had reinforced my behaviour after saying rules similar in function to "you can do this if you try"; this would therefore have strengthened a certain class of my pliance behaviour. It would be almost as if I tried simply *because* the rule was given to me by myself.

Unfortunately, pliance can be destructive when insensitivity to the natural contingencies is *undesirable*. According to Zettle and Hayes (1982), plies may often become overly rigid or insensitive in this way through the use of certain autoclitics. Recall, an autoclitic is a verbal speaker unit which alters the function of the accompanied verbalisation. Take the rule "I am not a failure provided I *always* get A's". A ply such as this may serve the long-term function of evoking consistent studying and the earning of A's, but as a functional autoclitic, the word *always* appears to change the function of the

statement from "I am not a failure if I get an A" to "I am a failure if I get any grade other than an A" (i.e., the function of the statement is now greatly altered). Thus, the potentially severe punishment of "feeling like a failure" if a "B" is awarded illustrates the dangerousness of plies that demand rigid adherence in order to prevent aversive consequences[11]. This type of RGB is labelled here as *overly rigid pliance.*

Cognitive therapy may often attend to overly rigid plies. Poppen (1989) points to Ellis' (1962) concept of *musturbation.* According to Ellis, people have a tendency to make harsh demands of themselves, such as "I *must* always succeed", and "I *should* be able to date a beautiful woman". Words such as *should* and *must* often change plies that once specified an attainable level of responding to rules which demand consistent or flawless standards of behaviour. Although such rules might be appropriate for continuous or consistent schedules, the extreme excess or deficit in behaviours they produce are often not required by the natural contingencies to which they point. Yet the discrepancy between the ply and the non-verbal schedule is not normally apparent to the rule-follower, because pliance, overly rigid or otherwise, has been learned through *speaker-mediated* reinforcement for rule-following rather than the identification of rules which correlate with the environment. Thus, plies are perhaps most problematic when they override a desired sensitivity to the natural contingency specified in the rule, and this, too, may constitute an aspect of dysfunctional cognition.

3.5 Errors in Rule Construction

I have now described three categories of rules which may be related to cognitive therapy: intraverbally designed self-tracks, inadvertently maintained rule following, and overly rigid pliance. Each of these rules are essentially problematic because of the way they are followed and their functional properties (i.e., the insensitivity caused by pliance is not destructive per se, it is only when sensitivity is desirable that plies cause problems in this way). In contrast, I will now elaborate on problems of rule-construction, where for some reason it is the produced *rule* itself which is clearly at fault. I will first introduce

[11] From an operant point of view, however, it is interesting to note that such statements up until the point at which a B was earned may have *ensured* the obtaining of A's and the avoidance of B's.

some basic units of rule-construction, and then describe some disorders of rule-formulation that might cause problematic behaviour.

Earlier I outlined Bandura's (1977a) concept of the *self-system.* The first part of the self-system involves self-evaluation, or the construction of self-efficacy beliefs. If the verbalisation of a belief can be substituted for the verbalisation of a rule, then self-efficacy beliefs might be considered to be the *verbalisation of a contingency regarding one's own behaviour.* Bandura (1977a) states that there are four sources of information used to construct self-efficacy beliefs. Specifically, one can construct a self-efficacy belief on the basis of (a) her own physiological responses (emotional arousal), (b) her own behaviour in comparison to other peoples (vicarious experience), (c) the way she performs in response to the environment (performance accomplishments), and (d) her verbal interactions with others (verbal persuasion). As such, four types of rules can also be identified, each defined by the formal class of organismic activity they describe (Table 3.5a). I have labeled these as *physiological* rules, *vicarious* rules, *motor* rules, and *verbal origin* rules, respectively.

Table 3.5a Formal units of self-rule construction.

Bandura's (1977a) source of self- efficacy.	**Class of organismic stimuli recognised as constituting a contingency.**	**Formally described type of rule-construction.**
Emotional Arousal	Private emotional and physiological sensations.	Physiological rules.
Performance Accomplishments.	Responding to a Contingency	Motor rules.
Verbal Persuasion.	Recognition of a verbally described contingency about one's behaviour.	Verbal Origin rules.
Vicarious Experience.	Observing someone else respond to a contingency.	Vicarious Rules.

In the present sense a physiological rule discriminates privately experienced biological movement or disruption as being part of a contingency. For example, the physiological rule "If I yell at my children my heart beats very fast and I can feel my blood pressure rise" might be constructed after an outburst of anger at one's children. A motor-rule is a self-rule derived from an observation of one's own performance. For example, a motor-rule such as "I can throw the ball a long way" could be formulated by someone after having observed herself throw a ball. Verbal origin rules are self-rules based on the verbal behaviour of others. For example, a verbal origin rule such as "I can throw the ball a long way" might be constructed after someone else had said "you pick it up with your hand, move your arm quickly, and release the ball". A vicarious rule is a self-rule applied to oneself after watching the behaviour of another. For example, the vicarious rule "I can throw the ball a long way" could be formulated after watching someone else throw a ball. Vicarious rules might be considered to be analogous to their non-verbal counterpart, imitation.

We are now ready to discuss several ways in which the faulty construction of rules can cause dysfunctional behaviour. A rule is constructed when a contingency acts as a discriminative stimulus for a verbal description of that contingency. Two different people may be confronted with the same contingency, but the way in which they construct a rule about that contingency depends entirely on their histories of reinforcement for rule-formulation. How might one identify what some of these discrepancies look like? Beck (1976) places great emphasis on "distortions in information-processing". However, this might easily be substituted for "distortions in rule-construction". Accordingly, Beck's seven identified logical distortions, *arbitrary inference, selective abstraction, overgeneralisation, magnification and minimisation, ignoring the positive, dichotomous thinking* and *personalisation* can be seen as directly corresponding to an error in rule-formulation (Table 3.5b).

First, a client may continually construct rules which discount or conflict with her prior experience, other rules in her rule-repertoire, or aspects of the current situation, but for some reason are still recited and used. I have labelled such rules *naïve rules*. Naïve rules correspond to what Beck (1976) calls *arbitrary inference*; a conclusion that ignores important information or has been arrived at in the face of conflicting data. Thus, any rule

which fails to account for all the relevant contingency-stimuli available can be termed naïve. A rule such as "I never do anything right" might be a naïve rule if it was used in spite of a notable conflict with another one of the client's rules, such as "I can play football very well". Naive rules also overlap with Beck's concept of *overgeneralisation*, making a conclusion about many situations on the basis of a few incidents. Naive rules are similar to Beck's overgeneralisation because in both, the client has had sufficient contact with the components of the contingency to allow most people to make reliable generalities about it, but still greatly exaggerates the number of contexts to which the rule is applicable. Second, a client may continually construct rules which are based on an isolated instance of a contingency. I have labelled such rules as *premature rules*. Unlike naïve rules which fail to incorporate available stimuli, premature rules are constructed before all of the aspects of the contingency have become available, or before enough examples of the contingency have been observed so that reliable generalities can be formulated (i.e., rules). Premature rules most closely correspond to what Beck calls *selective abstraction*, a conclusion or conceptualisation of a circumstance based on a specific isolated detail. For example, take the client who says "this relationship is never going to work, so there is no use in trying". This might be considered a premature rule if it was constructed in the infancy of a relationship. Premature rules also share some common ground with Beck's concept of overgeneralisation, because both involve specifying a set of stimulus conditions without proof of the rule's breadth of application.

Third, a client may continually construct rules which greatly overestimate the reinforcement value of the consequences *or* greatly underestimate the aversiveness of the consequences. I have labelled these rules as *over-anticipatory* rules. Over-anticipatory rules correspond to Beck's concept of *magnification*, or viewing something as far more significant than it objectively is. An example might be the rule "If I lose weight, everyone will like me". Such a rule might cause excessive dieting despite it being an inaccurate description of the reinforcers available for weight loss. Fourth, the reverse of over-anticipatory rules would be the construction of rules which greatly *under*estimate the reinforcement value of the consequences *or* *over*estimate the aversiveness of the consequences in relation to the contingency in question. I have labelled these rules *over-inhibitory* rules.

Table 3.5b Faulty rule construction: A translation of Beck's (1976) distortions in information-processing.

Beck's (1976) Primary Distortions	Name of Rule Error	Potential Associated Problems
Arbitrary inference & Over-generalisation	Naive Rules	The rule does not adequately use the client's wealth of experience and prior formulation of rules to her benefit, and can cause the use of rules which are entirely inappropriate to the situation.
Selective Abstraction & Over-generalisation	Premature Rules	Rule may not be usable in other situations, and requires the client to wait and observe additional aspects of the contingency before a reliable rule can be constructed about it.
Magnification	Over-Anticipatory Rules	Schedule insensitivity of the over-responding type. May induce many wasted or useless responses, and may incur negative consequences for excessive behaviour.
Minimisation	Over-Inhibited Rules	Schedule insensitivity of the under-responding type. May induce many wasted or useless responses, and may incur negative consequences for deficits in behaviour.
Ignoring the positive	Rule Absence	Contingency shaped behaviours are relied on too much, benefits of rules are not discovered.
Dichotomous thinking and all-or-nothing-thinking	Rule Immoderation	Causes insensitivity in both directions. Rules may fail to maximise the possible benefits available from the contingency, but can also encounter aversive consequences for a failure to accurately track a contingency.
Personalisation	Incorrectly Individualised Rules	May prevent desirable behaviours which the individual actually has the potential to enact, but could also stimulate behaviours which may not be appropriate for the individual, but only for others.

Over-inhibitory rules closely correspond to Beck's concept of minimisation, the error of viewing something positive as far less significant than it is. For example, take the employer who commends her fashion designer for an excellent set of blueprints. If the designer subsequently creates a rule such as "even if I try hard all I get is a few smiles here and there", then the reinforcing value of the employers praise may not have been

accurately reflected in the rule, and may cause an unproductive reluctance to put effort into future designs.

Fifth, a client may simply fail to make a rule when it would be worthwhile to do so (Hayes, Kohlenberg, & Melancon, 1989). I have labelled this error in rule construction as *rule-absence.* Rule absence is the radical behavioural parallel of Beck's concept of *ignoring the positive,* a failure to attend to the favourable aspects of one's day. For example, if someone is given a compliment about their writing skills, they may benefit from a rule such as "if I submit good articles, they may well get published", but simply fail to create such a rule. Sixth, a client may only create rules that specify an extreme excess in behaviour *or* an extreme deficit in behaviours in relation to the discriminated contingency. I have labelled such a tendency *rule immoderation,* characterised by a failure to construct rules that specify an appropriate "mid-way" in terms of effort, duration, intensity, or frequency of responding. Rule immoderation closely corresponds to Beck's concepts of *dichotomous thinking* and *all-or-nothing thinking,* both of which are a tendency to categorise experiences in one of two extremes, such as complete success or total failure. An example of rule immoderation might be the creation of the rule "I won't go to my exam tomorrow because I'll probably fail". Thus, this student would have completely withdrawn from an exam because a more moderated rule such as "I will try anyway, because I might pass" could not be constructed. The rule "I will try anyway, because I might pass" would have been moderated because it specified a potentially productive midway between complete withdrawal (e.g., not going to the exam at all) and complete excess (e.g., 24 hours of intense study for an exam on the following day).

Finally, some clients may show a tendency to construct rules designed to control their own behaviour when the contingency is only applicable to other people, or perhaps nobody. I have labelled such rules as *incorrectly-individualised* rules. Incorrectly-individualised rules closely correspond to Beck's concept of *personalisation,* attributing external events to oneself without evidence of a causal connection. An example might be the client who sees somebody experience rejection after asking someone out on a date, and subsequently creates the vicarious rule "If she can't get a date then there is obviously no way I will be able to". Such a contingency may not necessarily be applicable to the

client and therefore this rule would be highly dysfunctional. As with Beck's logical distortions, these rules are not exclusive of each other (e.g., it is possible that the same rule could be both premature *and* over-anticipatory).

3.6 General Cognitive Theories of Psychopathology as RGB.

Although I have discussed some potential causes of pathological behaviour in terms of RGB, it is likely that a theory of psychopathology based on RGB would need to be much more disorder-specific. The reasons for this are simple: a clinician's major strength lies in her ability to use diagnosis to draw on a priori knowledge of a case in order to implement a suitable treatment plan. Although a verbal functional analysis of a psychologically distressed client could use the fundamental concepts outlined here, assessment of specific disorders would probably be much easier if the nature and content of specific rule statements were identified for some of the major disorder categories. Accordingly, for a cognitive-behavioural practitioner to effectively use concepts of rule-governance, cognitive theories of psychopathology would need to be assimilated also.

Beck (1976) himself has emphasised the need for any theory of therapy to have a tenable theory of psychopathology, and most cognitive treatments are based on a corresponding cognitive model of the disorder. For example, a common cognitive theory of panic disorder holds that clients will often misinterpret insignificant bodily sensations as a sign of serious physiological disruption (Clark, 1986). Because this misinterpretation can in turn contribute to an *exacerbation* of the physical sensations themselves, a full blown panic attack becomes increasingly likely. Such a theory would be a central reference point to a cognitive therapy which focused on controlling the client's misinterpretation of bodily signs, or stated in terms of rule-formulation, the accurate construction of *physiological* rules. In a recent article, Beck (1993) has identified many cognitive distortions as being typical of the major diagnostic categories in the DSM IV (American Psychiatric Association, 1994). If such disorders truly are dependent on cognition then they might also be seen as characterised by certain rules. The table below contrasts several possible examples of disorder-specific rules with the cognitions that

Beck has suggested as common for those disorders (Table 3.6). Note that where Beck's (1993) examples are structurally defined, the rules suggested in the right column are intended to represent instances of broader functional classes. In reality, it could be possible that any of these rules could be verbalised in a range of structural forms with the same functional effect. I have used rules which explicitly specify contingencies so as to expose their typical function.

Table 3.6 Typical rules which characterise several DSM IV categories.

Diagnostic Category	**Beck's (1993) Suggested Typical Cognitions**	**Possible Problematic Rules.**
Drug Addiction	Beliefs such as "It's okay to have a smoke this one time" and "I can't stand my boredom without a fix" are typical.	Rules such as "If I take this drug my boredom will be relieved" and "If I stop taking drugs I will endure much pain" are examples.
Post Traumatic Stress Disorder (PTSD)	Beliefs such as "this (rape) proves I am just an object" and "I am worthless because I felt some pleasure" are characteristic.	Rules such as "I am dirty because I have been raped" and "If my boyfriend finds out that I was raped he will leave me" are typical.
Obsessive Compulsive Disorder	Beliefs such as "I must be crazy to have thoughts like this" and "It will be my fault if I don't do something about the (presumed) danger" are examples.	Rules such as "I must wash my hands many times or I will become ill" and "If I don't check the light switch again I may have left it on" are characteristic.
Family Dysfunction.	Conflicting values, such as the father believing "A child need's discipline" while the mother believes that "A child needs continuous care and love" are typical.	Parents may be following such rules as "If I do not discipline my children every day they will become soft" and "I must gain the support of my children over my spouse or I will be unable to control this family" might be examples.

Sex Offenders.	Beliefs such as "Sex with my daughter will be good for our relationship and help her to mature" is characteristic.	Rules such as "If I abuse this child and tell her not to tell anyone I will not get caught" are typical.
Depression	Beliefs such as "I am worthless" and "there is no future for me" and "everyone is too judgmental" are examples.	Rules such as "Even if I try to do something enjoyable I will be unable to find pleasure in it" (verbalising learned helplessness) and "Even if I treat that person with respect they will think poorly of me" are characteristic.
Schizophrenia	Belief's such as "If I hear voices somebody is trying to control my mind" and "Being mentally ill means I am helpless" are typical.	Rules such as "I need to avoid contact with people because if they find out that I suffer from delusions they will treat me like a crazy person", and "I cant do anything because someone is controlling my mind" are characteristic.
Bipolar Disorders/Manic Symptoms.	Beliefs such as "I have exceptional powers and should use them" are typical.	Rules such as "If I sell my house and invest in shares I will soon be a very rich person" are characteristic.
Eating Disorders.	Beliefs such as "I am ugly" and "Fat people are worthless" are examples.	Rules such as "If I lose weight my family will love me" and "If I do not vomit I will soon get fat and look ugly" are characteristic.

The potential number of rule patterns that could be identified as characteristic of specific disorders is undoubtedly quite substantial. Unfortunately, this may also mean that the identification of rules is so client-specific that constructing a taxonomy of "disorder-rules" is limited in its use. Of course, diagnostic categories are secondary to a

functional analysis of problematic behaviour, which are, from the viewpoint of the behaviourist, always the guiding strategy for any treatment plan. This would also dictate the decision to include any new factors in the clinical behavioural literature, regardless of whether they were cognitive processes typical to the disorder or not. Nevertheless, in the process of understanding the cause of symptoms as due to "particular cognitive structures" to the role of "particular types of rule-governance", the following three guidelines might be useful:

1. Identifying what usual portion of the behaviours of the disorder are rule-governed and what proportion are contingency-shaped. For example, is social anxiety generally more self-directed than contingency-shaped? Conversely, depressive disorders may be unrelated to rules, and may be characterised by a notable lack of self-direction. Of course, for many disorders an identification of the ratio of contingency to rule-governed behaviours will be specific to the individual, but there may be some which are *usually* verbally directed as opposed to contingency-shaped or vice versa. Equally important is the possibility that in some disorders, what appear to be rules are merely corollaries of overt behaviour, while in some disorders they might clearly play a causative role.

2. The extent that self-rules contribute to the disorder as opposed to given rules. For example, while some disorders may typically involve many intraverbally designed self-tracks, another client may be emitting dysfunctional behaviour because she has placed herself (or been placed) in an environment where poor rules are constantly presented. Furthermore, the control of self-rules as opposed to given rules might typically change over the development of some disorders. For example, rules about addictive behaviours may begin as self-administered instructions (e.g., "When I am bored like this, a shot of heroin will cheer me up") but the following of such rules may eventually lead the client into situations (e.g., in drug houses) where poor rules are frequently specified by others (e.g., peers saying things such as "go on, one more hit won't hurt you").

3. Certain disorders might be characterised by the types of dysfunctional pliance, tracking, and rule-formulation mentioned thus far. A clinician aware of these would, upon completing a functional analysis, be particularly aware of what types of verbal behaviour to look for and to try and remedy. Similarly, knowing which disorders tend to be followed because of pliance versus those which are more related to tracking could be useful. For example, it could be that some types of delusional thoughts influence behaviour not "simply because they have been said" (pliance), but because they appear to actually tact a non-verbal contingency, as in "If I kill my father then I can destroy his plans to humiliate me publicly" (tracking).

3.7 Cognitive Personality Theory: RGB as a Reciprocally-Determined System

The final issue I will discuss concerning case conceptualisation and assessment is in the area of personality. According to Beck (1976, 1993), all therapies must have a corresponding theory of personality. A theory of personality designed to cohere with a theory of therapy helps a clinician to understand the goals of treatment according to the total psychological makeup of the client, and enables a therapist to assess the behavioural traits which are likely to enhance or conflict with progress in treatment. Given the importance of personality theory in cognitive therapy, the question is raised as to whether any cognitively oriented personality constructs can be redefined within the framework of RGB, and whether any movement in this direction would result in clinical benefit.

This does not seem entirely impossible. Perhaps two of the most relevant personality dimensions to cognitive therapy are the concepts of *sociotropy* and *autonomy* (Beck, Epstein, & Harrison, 1983). Sociotropy refers to a tendency to depend on the appraisal of others to confirm one's beliefs. In particular, a need for closeness and social reassurance are characteristic of the sociotropic trait. Autonomy refers to a tendency to rely on one's own self-affirmation for the validation of beliefs. Autonomy is reflected by self-evaluation based on self-imposed criteria, goal setting, self-drive, determination and other similar themes. Beck et al. (1983) have argued that both sociotropy and autonomy may contribute to psychological distress. For example, take someone who relies on her

partner's praise to confirm beliefs of self-confidence. In this case, sociotropic traits might cause relationship troubles and depression to interact. Alternatively, take the person who evaluates her self-worth by using measures of job-efficiency as criterion. In this case, autonomic traits could cause a vulnerability to depression when important deadlines are not met. According to Beck and Weishaar (1995), the acute symptoms of depression and other mental disorders may be preceded by a major stressor which is particularly detrimental because of its interaction with personality type

It is unclear what specific functional behaviours these cognitive personality traits aim to explain, but at a speculative level (and consistent with an operant model), the autonomic and sociotropic traits could simply be a tendency to construct self-rules which evoke pliance and tracking, respectively. Because autonomic traits reflect a tendency to differentially reinforce oneself for maintaining standards in concordance with their own beliefs, they are probably self-plies when expressed as rules. They would be *self*-plies because the person generating the rule is also following it, and they would be *plies* because reinforcement for compliance would be mediated by the rule-giver (i.e., by oneself). For example, the self-ply "If I can't meet this deadline I am a complete failure" could specify a contingency relating "the meeting of a deadline" to the self-administered consequence of feeling like a failure. Such rules are likely to be powerful determinants of behaviour because of the important contingencies they specify, and thus a tendency to use them for this type of control would be tempting to people whose rule repertoires are dominated by *dysfunctional self-plies*.

Alternatively, because sociotropic traits reflect a tendency to rely on the reinforcement of others for the confirmation or denial of beliefs, they are probably self-tracks when they control behaviour as rules. They would be *self*-tracks because the person generating the rule would also be following it, and they would be *tracks* because people other than the rule-giver (i.e., oneself) would be administering reinforcement for the evoked RGB. For example, take the rule "I must meet this deadline if I am ever going to please my boss"[12]. This would be a self-track because it specifies a contingency relating "meeting a deadline" and the natural rewards of "pleasing the boss" that the rule-

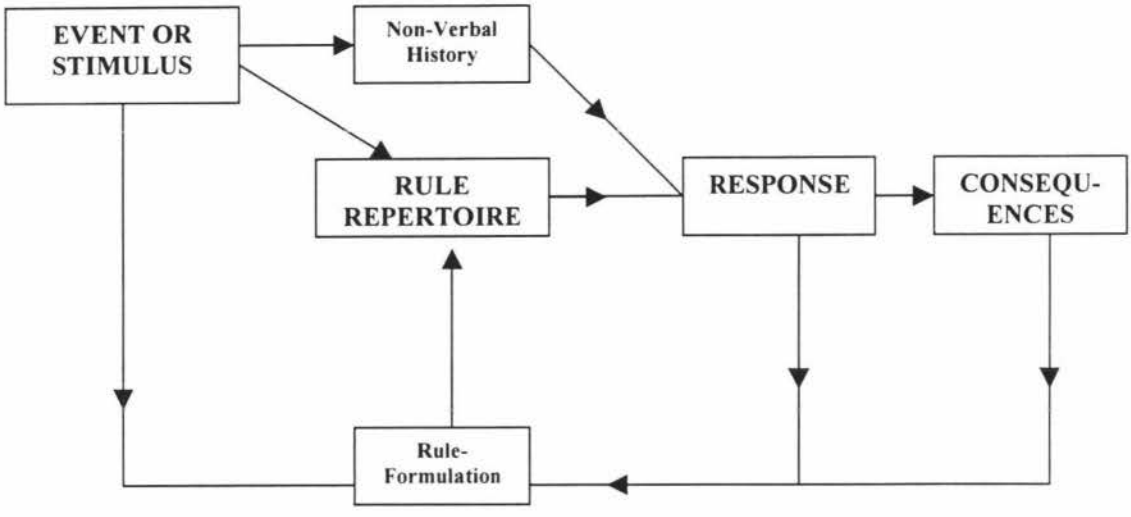
[12] As stressed in Part One, in the context of tracking, a natural contingency is considered *non-verbal* as it is independent of reinforcement mediated by the rule-giver.

giver (i.e., the employee) has no control over. If such self-tracks are relied on as a primary source of verbal control over one's life, it is understandable that a failure to meet the standards of the rule would produce negative emotional response. This may be a typical scenario for those individuals whose rule repertoires are dominated by *dysfunctional self-tracks*. However, clinically abnormal personality patterns would be addressed individually within the context of a verbal functional analysis, and do not seem particularly useful as a general guiding theme for a therapy based on RGB.

However, it is possible that Bandura's (1977a; 1977b; 1986) social-cognitive theory could assist in providing a general theory of human functioning that would be more compatible with the current emphasis on RGB. Some have argued that Bandura's theory is the most sophisticated, comprehensive, and integrated theory of human functioning that the cognitive–behavioural paradigm has to offer (e.g., Bergin & Garfield, 1994; Sweet & Loizeaux, 1991). As stated previously, central to Bandura's (1977a) self-efficacy theory is the assumption that people primarily control their own behaviour through *self-regulation*. According to Bandura, this occurs through a sequence of *self-observation* (examining one's own behaviour), followed by *self-evaluation* (judging the effectiveness of that behaviour), and finally *self-consequation* (establishing beliefs about the efficacy of oneself and the relevant behaviour). Furthermore, Bandura greatly emphasises reciprocal determinism, the contention that environment and cognition can not be separated, because self-determined behaviours influence the environment, which in turn influence cognition. Because Bandura's theory connects cognition and behaviour in this way, it is often viewed as a valuable integrative theory. However, a verbal analysis could just as well demonstrate how a person's rule-repertoire interacts with their behaviour, which in turn alters the environment, which in turn contributes to the creation of further self-rules (Fig. 3.7).

Figure 3.7 could be of benefit to the clinician trying to understand self-regulated behaviour from a radical behavioural perspective. The diagram shows the circular nature of the processes involved in self-governed rule-formulation and following. Any instance of operant behaviour begins with an event or stimulus. If this event has the potential to function as a rule, it will interact with particular components of the person's rule repertoire to form a response.

Figure 3.7 RGB as a reciprocally-determined system



If the event does not evoke the recitation of a rule, then it may only contribute to the response via the person's non-verbal history, or stated somewhat differently, learning experiences[13]. The ensuing behavioural event could then be used to construct more rules. Specifically, the response itself, as well as the consequences and the original context, can be discriminated against as the three elements required to construct a rule, which could then be added to the person's total rule-repertoire. A rule repertoire will be defined here as "the sum total of all of the rules a person has the potential to generate given the necessary stimulus conditions". A clinician may trace the cause of dysfunctional behaviour to one or more stages of the model: Is it the events themselves which are causing problems, or the rules they evoke? Is the client's rule repertoire causing a formal-functional error in translation? Could the problematic behaviour be devoid of rules and simply be interacting with past troubling experiences? Is rule-formulation faulty? If so, which aspect of the rule is poorly specified - the event, the behaviour, or the consequences? Viewing psychological distress as a systematic cycle of verbal and non-

[13] This is significant because it should never be *assumed* that a client's verbal behaviour is controlling other behaviours – such verbalisation may only be a retrospective report or an internal corollary of overt action.

verbal responding may help a clinician understand the way a client's verbal history is interacting with ongoing life events, without reference to the mentalistic terms that cognitive theories demand.

An example of employing this model might resemble the following. A man enters a bar and after sitting for some time, approaches a woman (Ms. X) and asks if he can sit and talk with her. She accepts the offer, and they begin to chat. Unfortunately, they do not get along at all, and Ms. X says several things that are quite offensive and derogatory to the man. This is a very aversive consequence to this man, and because he recognises the event as containing the three elements necessary in order to construct a rule, he describes the relationship between the stimulus (seeing Ms. X in a bar), the ensuing behaviour (asking her to talk and then engaging in conversation) and the consequence (embarrassment and rejection). Thus, the three elements constituting the contingency each contributed to the formulation of a rule, which might look like "when I see Ms. X, I should not try and talk to her because she will reject me". On a second occasion, the same man sees Ms. X again, but this time, as a stimulus, he recognises her as associated with the rule, and thus covertly recites and follows the previously formulated "rejection by Ms. X rule". Thus, the situation (seeing Ms. X) has interacted with a new component of his rule repertoire (Fig. 3.7)

However, rule formulation had worked to the benefit of this man on *this* occasion - perhaps what some would refer to as serving a type of "defense-mechanism". If he had instead constructed the rule "when I see a *woman*, I should not approach her because I will probably get rejected", such a rule might be used in hundreds of other situations in which this man was in the presence of women, and he may even construct further rules which lead to an avoidance of women in general (i.e., reciprocally-determining his environment). However, by seeing RGB as a different type of self-system, we can try and trace the cause of the dysfunctional rule to one or more specific points. In the second instance of the example, the stimulus in the rule was poorly described. This man only had sufficient experience to construct a rule about Ms. X, not *women* in general (i.e., it was a naive rule, or what Beck (1976) would refer to as *overgeneralisation* or a *selective-abstraction*). It may also be the case that this man simply did not tact all of the behaviours of his which caused Ms. X to "reject" him or turn him away. If this was the

case, then the behaviours described in formulating the rule may have been unrelated to the programmed contingency (e.g., a more accurate rule may have been "If I don't swear repeatedly, I will have a greater chance of obtaining Ms. X's approval). It may be the case that this man needs first to be taught how to socially relate in an effective way. Thus, a secondary function of seeing RGB as a reciprocally-determined system (Fig. 3.7) is that it can point to areas needing further assessment or *verbal-probing* before an intervention can be effectively conducted.

3.8 Cognitive Mechanisms of Change as RGB.

I have now offered several ways in which a clinician might assess dysfunctional rule-following and formulation from a radical behavioural perspective. Of course, this would have to be complemented with a functional analysis of non-verbal deficits and/or excesses also, but I will discuss the ways in which RGB and contingency management can work in unison later. Next, however, I will address the fact that there must also be a tenable theory of the process of change in any system of psychotherapy (Beck, 1976). I will discuss the major ways in which rules could be effectively used as a method of behavioural change, and discuss how these methods might be seen as directly related to the mechanisms of change on which cognitive therapies rely. I have divided these "rule-techniques" into three main areas: (a) directly eliminating the presentation of dysfunctional rules, (b) training clients in appropriate rule-formulation, and (c) directly giving a client rules to follow. Each of these can be seen as separate techniques in theory, but could obviously be combined in application (e.g., amending a long-held dysfunctional rule would essentially involve both extinguishing its presentation *and* formulating a new one).

1. Punishing or Extinguishing the Recital of Poor Rules. The most obvious way to prevent the following of dysfunctional rules would be to directly eliminate the presentation of the rule itself. That is, a therapist can essentially teach a client to recognise the presentation of problematic rules and then eliminate this behaviour of the

speaker (whether this is the client's behaviour or someone else's). This is related to my earlier discussion of inadvertently maintained dysfunctional rules. Inadvertently maintained rules are reinforced for reasons other than their function as precurrent behaviour. Conversely, I am now arguing that obviating poor rule-recital itself can be used to indirectly eliminate problematic RGB (by indirect, I mean in contrast to directly punishing the rule-induced *behaviour*). Earlier, I used the non-verbal example of a therapist eliminating the kitchen area as a source of problematic stimulus control for undesirable eating behaviours. Here I am simply suggesting that although rules are verbal stimuli they are usually *behaviours* as well, and therefore may need to be eliminated in order to undermine the problematic RGB they stimulate - this would be akin to administering punishment for entering the kitchen areas at times outside of designated meal periods. Presumably, the behaviour of rule-recital can be eliminated in the same way that any other behaviour can, by consequating it with an aversive agent (punishment) or by removing any positive agents that may be reinforcing the presentation of the rule (extinction).

In some cases, a client's exposure to faulty rules may primarily occur via external speakers. As such, external sources need to be examined when considering the elimination of faulty rule-recital. A client could be trained to punish the presentation of external speaker rules herself. For example, a person who is constantly berated by her boss with rules such as "you are useless and can't do anything right!" may wish to punish this behaviour by threatening to file a harassment complaint. It is also possible that in some circumstances the *therapist* may punish rules given to the client by an external speaker. For example, in the context of family therapy, it may become obvious that the father frequently issues punitive rules to the child (client) such as "be quiet, so we can solve this problem!". A therapist may punish this behaviour by pointing out the problems that such rules cause (e.g., the drastically lowered probability that the child will cooperate).

However, cognitive therapists probably focus more on eliminating the recitation of dysfunctional *self*-rules, often through punishment or extinction initiated by either the therapist or speaker. It is probably only when the therapist can evoke an overt verbalisation of poor self-rules that the therapist can directly punish such behaviour. For

example, the cognitive therapist might say "what are the automatic thoughts that are coming into your head, now we have decided to discuss your phobia of heights?", and the client might then say "I know I'm going to get nervous if we talk about heights". In an attempt to eliminate the recital of such a self-defeating rule, the therapist might then respond with a mildly aversive statement such as "well, if you *say* you're going to get nervous, you probably will". Of course, the therapist is only physically present for a short time in the client's life, so training clients to develop strategies of punishing their own problematic precurrent rules would be a more enduring method of change. This would involve training the client to recognise particularly bad self-rules which have been in operation for some time (which may be many of the problematic rules I have already discussed). For example, a therapist may notice that self-defeating statements such as "I need to call in sick today because I know I am going to screw up that speech" tend to function as rules which evoke avoidance behaviour. Thus, the client may greatly benefit from recognising such behaviour and punishing it. In this example, the client may say to herself "If tell myself that I'm going to screw it up, I probably will", or "there I go again, *telling* myself that I will fail when I will probably do alright!".

Directly eliminating the recital of undesirable rules may be related to Beck's *guided discovery,* where the client comes to realise that many past situations can be seen as a result of similar cognitive distortions (Beck & Weishaar, 1995). If such "cognitive distortions" can be simply seen as problematic self-talk, then perhaps guided discovery may also train the client to recognise and punish the recital of long-held classes of dysfunctional self-rules. Beck (1976) uses what he calls a *daily record of dysfunctional thoughts,* so that the client can monitor "changes in thinking" and other problematic self-talk. Such self-monitoring procedures may work because they allow the explicit recognition of poor rules, and allow the client to immediately punish her own rule-recital (e.g., "now look what I just said to myself!") and/or stimulate the client to introduce a rule incompatible with the problematic one (e.g., "what could be a plausible alternative to that thought?").

Alternatively, Ellis (1962) prefers to refer to directly punishing the presentation of poor rules as "attention to the four D's", where irrational beliefs are immediately *detected, disputed, discriminated* and *defined.* Although colourful descriptions, Ellis' four

D's are probably just terms for describing the evocation of long-term dysfunctional beliefs, which could be functioning as rules. Upon evoking such rules, Ellis would generally punish their recital, sometimes with confrontational exclamations such as "No! that's just irrational!". Ellis (1962) typically continues to repeat this process until the client responds with alternative beliefs, seemingly independent of outside encouragement. Unfortunately, with a therapy as confrontational as REBT, there is the risk that some clients will offer "rational" beliefs purely in order to please the therapist or to avoid her rebukes (Zettle & Hayes, 1982). That is, in some cases a client might appear as though she is formulating a rule, but is simply exhibiting loose imitations of the therapist's verbalisations in order to please the therapist. An alternative to this, and as will be advocated under the next heading of training rule-formulation, might involve a collaborative shaping of the formulation of rules which are more adaptive. This would ensure that the client actually learns to *identify* problematic rules, rather than to simply submit to the REBT therapist's requests to abandon poor rules, or to "admit" that they were mistaken.

Beck (1976) has suggested that depressed individuals commonly make upsetting statements to themselves that both undermine their confidence and inhibit adaptive behaviours. According to Beck, such thoughts can be *automatic* in that they occur almost unconsciously and without forethought or contemplation of the relevant situation. Now consider the fact that if the verbal expression of rules is operant behaviour, they are probably often shaped into abbreviated forms which make it extremely hard to topographically identify any implicit contingency (Jaremko, 1987). As precurrent behaviour, self-rule presentation becomes more subtle as formal qualities are omitted because of their redundancy. For example, a person with a social phobia may, when in the presence of people, covertly recite the rule "People do not like me once they start talking to me", and such a rule may persistently induce social withdrawal. The more the rule is used, the more refined it may become until just remnants of it are left - perhaps just private utterances such as "I should not socialise" or "don't talk" and other brief comments which serve to protect the individual with social avoidance behaviour. A rule may even be reduced to a simple associated image; applied to the current example, the image of a previously embarrassing social situation could induce the RGB of social

withdrawal. Identifying "automatic thoughts" may simply be the process of prompting the client to verbalise the function of such abbreviated rules at *the time of the rule-evoking situation*, usually by asking them to report the conditions and emotional responses that have surrounded them in the past. As the client comes to "catch" these thoughts, they are better able to punish their presentation, or construct and recite incompatible rules (e.g., "that embarassing situation was a one-off, I know I can enjoy the company of these people").

So far I have only spoken about *punishing* rule-recital in order to undermine the ensuing undesirable RGB. However, the reader may note that in the initial paragraph and title of this subsection I also explicitly referred to the use of *extinction* as a means of eliminating problematic rule-recital. There is a reason for this extra emphasis on *punishment* rather than simply removing any extraneous reinforcement that may be maintaining rule-recital. Recall that in the introduction I explained how presenting oneself with a rule is precurrent behaviour because it serves to generate a discriminative stimulus for potentially reinforceable behaviour. In other words, the discriminative stimulus, or rule, acts as the end reinforcement for rule-formulation. One cannot overlook the fact that this is an important source of reinforcement. People would probably not allocate time to the "thinking" of a rule if such formulation did not usually result in the generation of a rule. Thus, any other reinforcement for rule-recital (e.g., anxiety relief, agreement from others) is probably secondary to this function in most cases[14]. Thus, in order to compete with the reinforcing effects of producing useful discriminative stimuli, punishment, rather than extinction, is probably required in most circumstances.

However, removing those factors which inadvertently maintain rule-following may mean that *less* punishment of rule-recital-behaviour is required for its elimination, and this might produce a less intrusive or effortful form of therapy for the client. Furthermore, if a therapist can determine the collateral contingencies operating on the recital of a particular class of undesirable rules (i.e., in addition to their function of inducing RGB), then the therapist may also be able to identify basic needs in the client that will be left

[14] The exception to this is when other contingencies for rule-following *override* the negative effects of producing a misguided rule. Hence, my earlier recommendation of assessing for *inadvertently-maintained rule-following* when conducting a functional verbal analysis of a psychologically disturbed client.

unmet once the rule has been eliminated. The value of using extinction to diminish dysfunctional rule presentation in this manner becomes clearer when one views the use of this method in cognitive therapy. Mackay and Fanning (1987) have written extensively about the role of negative self-talk, and have placed these ideas within their cognitive self-help program for maintaining and improving self-esteem. They claim that people with poor self-esteem possess an "inner-critic" which constantly tells them facts and assumptions about themselves, but in an irrational and distorted way (they use the term *critic* to allow the client to externalise the source of their negative self-talk);

> You listen to the critic because it is rewarding to do so. Incredible as it seems, the critic helps you to meet certain basic needs, and listening to his vicious attacks can be reinforcing...When the critic helps you to diminish or entirely stop painful feelings, his voice is highly reinforced. Even though the long-term effect is to destroy your self-esteem, the short-term effect of critical self-talk may be a reduction in affect...It's extremely important that you learn to identify the function of your self-attacks, how they help as well as hurt you. But knowing that function doesn't change much. Those same needs must be met in new and healthy ways (McKay & Fanning, 1987, p. 21).

McKay and Fanning (1987) are implicitly referring to the idea that in addition to the role self-talk can play in causing RGB, it may serve other, ancillary functions. Later in the book, they seem to be suggesting that therapists must introduce other methods to compensate for these functions if they are to remove this source of reinforcement for negative self-talk (i.e., before it can be extinguished). As an example, McKay and Fanning point to the role of self-talk in dealing with anger:

> Feelings of anger toward people you love can be very frightening. As the anger begins to enter awareness, you may feel a huge surge of anxiety. One way of coping is to turn anger around and attack yourself. You're the one who's failed, who hasn't understood, who's mistakes cause the problem in the first place. As the critic goes on attack, your anxiety decreases...The critic helps you deal with your fear of anger by deflecting into an attack on yourself. A healthier strategy for dealing

> with your anger is to learn to say what you want and negotiate for change. Anger is so often a byproduct of helplessness because your needs are expressed or unexpressed ineffectually (p. 39).

McKay and Fanning (1987) are arguing that in order to remove the negative self-talk that results from anger, a person needs to both acknowledge and compensate for the basic needs that such self-talk fulfills. This argument might be greatly clarified by using the concept of RGB. From a radical behavioural perspective, anger is probably most detrimental when it causes a person to instruct herself to do things which are extremely damaging to the individuals well-being or self-esteem. For example, take the violent prison inmate man who becomes extremely angry over a conflict with another, and says "I'm going to hit that guy and teach him a lesson ". If he was to follow this rule, his RGB has probably caused serious detriment to his own situation and that of the person he is in conflict with. Teaching the client to punish such a rule upon recital may be an immediate way to reduce its likelihood - this client may be taught, after the recital of such a rule, to say "when I'm angry and say things like that, I do them without any clear reason!". However, as it relates to extinction, it may also be wise to withdraw collateral reinforcement for such rules. For example, the man in this example may have found that the verbalisation of such "tough-guy" rules resulted in admiration and commendation from peers who also engaged in such behaviour. Thus, dealing with an anger problem such as this might first involve changing the client's social situation so that the reinforcing behaviour of these peers were no longer a part of the client's environment (e.g., moving him to a quieter or less-violent prison unit). Alternately, we could ask the client himself to arrange the removal of such collateral reinforcement, perhaps by saying to the other inmates "Listen you guys, don't encourage me when I get worked up like that because I always seem to end up in solitary confinement". Thus, by removing all other sources of reinforcement for the verbalisation of dysfunctional self-rules, a therapist may be in a better position to reduce the verbalisation of damaging self-rules, whether by punishment or by other means.

However, the consistent recognition of faulty rules and the cessation of their use would probably result in minimal effects on its own. As verbal beings, clients would need

to learn how to *make* new rules to fill the vacuum left by old rules. As such, I will now discuss the training of *rule-formulation* as a method of behaviour change.

2. Train the behaviour of formulating useful rules. Effective rule-formulation could be a useful behaviour for people with psychological problems to learn, and may constitute an integral part of cognitive therapy. I will discuss training the formulation of both plies and tracks.

Most people probably use self-plies as a way to override prevailing negative consequences. That is, when someone needs to maintain behaviour which only provides small, distant or only cumulatively significant reinforcement, a ply can provide verbally-mediated control which will suffice until the long-term gains provided by the RGB's are earned. Thus, plies are insensitive for good reason, but as discussed earlier, problems occur when plies demand too much - at least, much more than what is required by the contingency in question (i.e., when they are overly rigid plies). For example, someone may be responding to a contingency of physical labour. In order to override the punishing effects of physical exertion, someone might say to themselves "I need to finish all of the gardening *today*". Now assume that this person was being paid to finish this job on a contractual basis, and the time limit for the work was not relevant. Thus, even if the physical labour was spread out over a week, the same amount of money would be paid, yet such a rule would enforce extremely difficult and labour-intensive behaviours for no extra gain (assuming this person had ample time on her hands, and was not trying to avoid the punishment of having less free time). Thus, the rule would have been overly rigid and highly insensitive to the natural contingency its evoked behaviours may be in contact with. Training in ply-formulation would teach the client to create rules which specify behaviours which are both *sufficient* and *efficient*, and do not cause behavioural excesses or deficits.

If a client is to learn to use plies to her advantage, then presumably conditions which maximise the likelihood of pliance should be employed (Zettle & Hayes, 1982). Recall, the likelihood of pliance is a function of the apparent ability of the speaker to administer reinforcement. Thus, pliance is often dependant on the speaker's ability to monitor compliance. For example, self-plies which specify hard-to-monitor behaviours (e.g., "I

must think about my positive aspects when I feel ugly, and this will help lift my depression") are likely to be more effective if the client can record these behaviours, and thus differentially self-reinforce varying approximations of compliance (e.g., "good, I thought of three positive things about my appearance!"). If a client is to be trained how to use *externally*-mediated self-plies, then the speaker needs to ensure that the listener is conspicuously prepared to administer consequences, as this also maximises the likelihood of pliance (Zettle & Hayes, 1982). For example, a therapist may help a disruptive child to create self-plies such as "Mom, I promise I will do the lawns everyday for the next three weeks". Mum, who has already paid commercial gardeners in advance to mow the lawns for the next month would probably be indifferent to her son's lawn-mowing, and would be unlikely to arrange consequences to reinforce compliance. If this was apparent to the child, the rule may not function as an effective self-ply.

Similarly, if a therapist intends to teach someone how to learn to reinforce *themselves* for pliance, then they need to ensure that the client arranges that reinforcement (self-reinforcement). For example, if someone constructs a ply such as "If I work until midnight, I can take an extra day off next week", the salience of their ability to carry out such reinforcement will increase the likelihood of compliance (e.g., perhaps this person really knew that it was unlikely that a day off the following week could be arranged, and thus pliance would be less likely). Of course, the likelihood of pliance is also dependant on how important the listener-mediated reinforcement is to the speaker. The availability of a day off work would hardly ensure compliance if this person was a "workaholic", in which case, time off work would be more likely to be aversive. Thus, when training the formulation of plies, a therapist should be aware that (a) there are factors which maximise the likelihood of pliance, and (b) clients can be trained to use the schedule insensitivity of pliance to their maximal advantage, but (c) clients need to use plies that produce behaviours which do not exceed what is required by the natural contingency providing reinforcement.

Cognitive therapy seems to rely more on teaching the behaviour of constructing "realistic" or "rational" rules, or rules which are in some way connected with the actual state of affairs - the same rule-environment correspondences which determines the effectiveness of *tracks*. Using the concept of RGB, this might first involve the therapist

training clients to tact contingencies in ways that are testable, or discriminating between contingencies that the client would benefit from describing as a rule from those that she would not. This would probably occur by the therapist *shaping* track-formulation, or perhaps having the therapist *model* track-formulation. How might cognitive therapies do this? One way could be through the previously described *Socratic method of questioning*. The Socratic method of questioning is a core part of cognitive therapies, most notably Beck's CT (1976). According to Beck and Weishaar (1995), the Socratic dialogue is designed to gently probe the client's thought patterns using questions that can (a) clarify problems, (b) identify and assess thought patterns, (c) examine the meaning of events to clients, and (d) illustrate the consequences of maintaining certain beliefs. Thus, a central function of Socratic questioning from a cognitive perspective is to convey information by guiding the client's natural train of information-processing (Beck, 1976). However, its function may be described more precisely by taking an operant approach.

The term Socratic *questioning* could mistakenly lead one to assume that its use is primarily that of an assessment tool. While the Socratic dialogue probably does play a role in the discrimination of problematic rules, for it to have therapeutic impact it must also *teach* new verbal behaviours. Thus, from a radical behavioural perspective, the Socratic method of questioning might be viewed as an extended verbal episode designed to aid the training of rule-formulation. This may very well be seen as an oversimplification if one were to ignore the fact that a unique feature of the Socratic dialogue is that while it appears to shape and model verbal behaviour (because the therapist is only *questioning* the client), it seems to give her the *impression* that such training is actually *self-discovered*. It is hard to conceptualise this within behavioural theory, and it is unclear why Beck (1976) has nominated the Socratic dialogue as the "preferred" method of therapeutic discourse, but the findings of the behavioural experiment outlined in the introduction by Catania, Matthews and Shimoff (1982) may provide a provisional explanation.

Recall that the Catania et al. (1982) experiment showed that although training rule-formulation required prolonged periods of shaping before the produced rules effectively evoked behaviour, once established, those rules controlled behaviour more reliably than rules that had been simply presented outright. Extrapolating these findings to a clinical

setting, it is possible that the rules generated by the Socratic dialogue could be more effective in evoking RGB because the client's self-rules have been *shaped* by the therapist, as opposed to having been given outright. Thus, it may be the case that self-rules created using the Socratic dialogue are a better way to reliably ensure that those newly learned rules will reliably control behaviour in situations outside of the therapeutic session. Furthermore, the Socratic dialogue may inadvertently train useful rule-formulation through *modelling* patterns of rule-construction. By observing the chain of questions that the therapist asks, a client can watch firsthand how a person who is not in psychological distress creates adaptive rules. Even thought it is the client's responsibility to independently demonstrate her ability to produce adaptive rules in the later stages of therapy, the Socratic dialogue may greatly assist a therapist in *modelling* rule-formulation in the initial stages. Stated somewhat differently, a client may have learned "what questions they should ask themselves" when engaging in rule-formulation, having observed many times the questions what a therapist would ask.

Another important aspect of track formulation would be the actual *testing* of tracks. That is, following a track and then re-evaluating its correspondence with the natural contingency it specified. For example, an unassertive architect may say "If I show creativity at work I will be frowned upon by my boss". This is testable, and a client may follow it to assess the correspondence between the rule and the state of affairs at work. If the rule is found to be true and accurate, it may be desirable to simply refrain from showing creativity at work. However, following it without constructing a rule which can test this proposition might maintain a very monotonous work schedule, for no tangible benefit. For example, having been directed by the therapist to try an alternative method of responding to this contingency, it might be found that by showing a *small* amount of creativity some approval could be earned. At this point, the client could construct the track "As long as I don't go overboard with my designs, I can immerse ideas of my own without being admonished by my employer". Thus, testing tracks would shape track-formulation, because if a self-track is found to be inaccurate then that same track or its properties would be less likely to be incorporated into the formulation of future tracks. Of course, in order to allow a client to *practice* new rule-formulation skills, or to gain contact with new contingencies for which a rule could be profitably formulated,

homework exercises which led the client to these contingencies would be necessary. This also seems to be a primary mechanism by which cognitive therapies train track-formulation.

In fact, both Ellis (1995) and Beck (1976) strongly advocate the use of *hypothesis-testing* and other exercises which are designed to modify beliefs. However, rather than simply confirming beliefs as "plausible" or as more "realistic" interpretations, such exercises are probably just allowing the formulation of new and improved tracks. Although Ellis (1962) argues that new "rational" beliefs are simply a product of discovering reality, if such beliefs are only significant in as far as they control behaviour (i.e., as far as they are rules) then they may be more accurately seen as a client "discovering" the discrepancy between their old self-tracks and the environment, and subsequently be encouraged to construct a more plausible track. Ultimately, the therapist aims to teach the client to schedule her own homework exercises, so that she can independently come to use direct experience as a means of shaping and maintaining the formulation of new and more adaptive tracks. This could be the behavioural change that Beck (1976) refers to as the *cognitive shift*, the process whereby clients come to re-evaluate day-to-day situations in a more "realistic" and adaptive way. Thus, when training track-formulation, therapists should ask themselves "is the client tacting the most appropriate contingencies, and if so, are these rules an accurate representation of those contingencies?"

One advantage of taking a radical behavioural approach to CBT in this way would involve refraining from offering intraverbally designed rules as a replacement for problematic rules. This problem seems to be most often associated with REBT (Zettle & Hayes, 1982). For example, an irrational rule might be "The world is an awful place". It is possible that an REBT therapist could train the client to verbalise new "rational" rules, such as "the world is a great place". However, as outlined earlier, words such as *great* are intraverbal verbalisations rather than a direct tact of a contingency and its relevant stimuli, and are thus untestable as tracks (i.e., there is no specific behaviour one could engage in which would confirm or deny the accuracy of "the world is a great place"). Although new tracks which involve intraverbals in this way may be temporarily pleasing to the client, they would not be the best rules to assist in long-term behaviour change as

the testing of tracks is an essential part of their successful establishment within the client's verbal repertoire. Thus, intraverbally designed tracks, no matter how "rational", would probably fail to evoke behaviours that would persist over time.

Finally, the reader may note that when giving examples of training rules within the context of an RGB C-B integration, I have only used rules which very specifically describe certain contingencies. This is simply because training the formulation of rules which are specific in regards to the contingencies they point to is viewed here as a potential improvement on *all* cognitive and semantically-oriented therapies. *Because cognitive therapies primarily place emphasis on meaning and the emotional effects of certain self-talk, they may overlook the possibility that beliefs which specifically and clearly relate to new behaviours could be the most reliable way to verbally alter behaviour.* In contrast, and as a direct improvement on the cognitive model, RGB C-B approaches would greatly emphasise the importance of shaping and training performance-specific rules that explicitly contained the three components of a contingency, especially the *required* behaviours. The empirical study outlined in Part One by Matthews, Catania, and Shimoff (1985) would support this use of performance-specific rules, as it was shown in this experiment that such rules exert more control over behaviour than general "contingency statements".

Such specificity may be also be very important as it relates to rules which clearly define those situations for which the rule is most appropriate. Beliefs acquired throughout a *cognitive shift* (Beck, 1976), or a general emphasis on "rationality" (Ellis, 1962), could be relatively oblivious to the contexts for which they are most suited (Poppen, 1989). As such, it is easy to picture how rationality could be *overdone*, such as circumstances where behaviours *need* to be punished so they do not occur again (e.g., acting rudely and then "rationalising" the consequences). In addition, even rules which are clearly adaptive but do not clearly specify a context might cause a type of generalised punishment for following other therapy rules. For example, a client who excessively "rationalises" situations or follows rules blindly may incur aversive consequences for doing so, and such punishment may cause the client to identify *all* therapy-learnt rules as being useless or misguiding. However, an RGB C-B approach fully acknowledges that therapeutic rules are often appropriately bound to certain situations, and thus all rules formed in

therapy would be constructed in a way that does not lead to an overgeneralisation of their use. Of course, the therapist would have the most control over the specificity of therapeutic rules when offering pre-formulated rules to a client, the process of which I will explain now.

3. Give clients rules to follow outright, and teach them to recognise situations where those rules should be stated and followed. A therapist may simply decide to give the client rules by simply telling them what to do, that is, exposing the client to a specific rule which can then be recited later. For a therapist to give a client a track to follow, without appearing as though she will monitor and reinforce compliance, would closely resemble giving "advice". However, in most situations, therapists would presumably instruct speakers to recite a specific self-rule in a given circumstance, and then consequate a correspondence between the rule and the behaviour it evokes. In order to do this, the therapist as speaker would usually monitor compliance by reviewing third-party reports (e.g., behavioural checklists, observational reports, etc), or question the client in a way that evokes a retrospective report of the level of abidance with the rule (e.g., "did you breath slowly when you saw your ex-husband, as we agreed you would instruct yourself to do?"). This, of course, is the type of socially-mediated control that defines pliance.

The direct administration of plies by a therapist might be used to change behaviour in four main ways. The first two relate to in-session behaviours. First, a ply can be given by the therapist, followed by the client, and reinforced by the therapist for accurate compliance. For example, a rule such as "ok, breathe deeply, relax your muscles, and converse with me like any other person, and I will talk with you" can be both stated and reinforced by the therapist (in this case, reinforced with conversation). Second, the same ply could be recited by the client, but still reinforced by the therapist. For example, take the rule "If I breathe deeply and relax, I can practice socialising effectively with you". If the behaviours to be induced with given rules can not be practiced in therapy (e.g., masturbating as part of sex therapy), the therapist has two further options. First, a rule can be recited by the client but in a way that specifies self-reinforcement. An explicit self-ply such as this might look like the covert statement "because I will not think of my sexual

unattractiveness I will be able to maintain my arousal, and afterwards I will give myself credit for being a great lover". Second, the client may be trained to recite rules as plies in ways that make a social commitment, so that certain other listeners can ensure compliance (Skinner, 1969). For example, a therapist who has arranged an exercise program for someone suffering from obesity might instruct them to tell their spouse about their homework assignments, such as "this week I am going to run every morning – you just see if I don't!". Such a statement would act as a ply because the client's spouse would presumably punish a failure to comply with the self-imposed task.

As rule recital is behaviour, a therapist should ensure that such behaviours are easily generalised to settings outside the laboratory by using contexts of rule-rehearsal that are both diverse and accurate. For example, a therapist may wish to train the client to covertly recite a given rule in a particular circumstance, especially if that given circumstance is likely to be at a time when anxiety or other matters could potentially divert attention from saying the rule. For example, a therapist in session might say to a client suffering from an anxiety-related sexual disorder: "now imagine you're about to have sex and you're feeling anxious, practice breathing deeply and calmly, and say to yourself.... ".

Finally, when the emphasis is on training the client to recite the given rule in real-life circumstances, teaching the client to recognise the exact situations in which the rule will be useful will prevent an over-generalised usage of that rule. For example, telling a psychopathic client to always use the rule "people appreciate humour" may be problematic if the client was not also told also that some humour can be offensive in some circumstances (e.g., funerals). Training contextually bound rules would simply involve teaching the client to be specific about the circumstances for which their self-constructed rules have been designed. Finally, some rules might simply be given to the client so as to compete with other, less helpful rules. For example, take the person who says to themselves before giving a speech; "I need to get out of here because I *know* I'll screw it up when someone asks me a question". This person might benefit by also saying; "Even if I can't answer the question it will not affect my academic standing". Thus, if a clinician is unable to extinguish faulty rule-recital and following, administering a competing rule may be a viable option.

Of course, directly giving context and behaviour-specific instructions to a client, which can then be used to induce adaptive behaviour, is most reminiscent of SIT (Meichenbaum, 1985). Zettle and Hayes (1982) point out that it has been empirically shown that the use of coping self-statements in SIT does not suggest that their function is to relieve or divert attention from aversive states of anxiety, but more likely to promote an increased sense of mastery in problematic situations (Kanfer, Karoly, & Newman, 1975). As such, Zettle and Hayes argue that attempting to gain such mastery with the following of covertly-recited coping statements may occur through learning to identify problematic situations for which an identified rule, originally given by the therapist, should evoke an adaptive response. Because the nature of these statements are competence related, they are most probably functioning as plies which are verbally reinforced by either the therapist or client. That is, because the contingencies pointed to by typical SIT rules are usually unfamiliar to the client, SIT most probably works through ply administration. For example, a client would not normally track a relationship between breathing deeply and successful social interaction if they had never before been reinforced for following such a rule (or similar rules), whereas a ply would not depend on any such correspondence between the rule and this environmental contingency of social engagement.

Viewing the SIT therapist as a ply-giver would imply that directly telling the client to engage in certain self-instruction is likely to be effective if compliance was self-reinforced, such as the client saying to herself; "I knew I could talk more comfortably if I followed the deep-breathing rule that my therapist taught me". However, in most cases the therapist would provide reinforcement for a correspondence between the given rule and the evoked behaviour (e.g., verbal praise such as "I think you are showing improvement because you recited the statements I trained you to, and followed them accurately"). As explained in Part One, self-pliance can occur through reinforcement given by an external listener, because to qualify as pliance, RGB only has to be controlled by reinforcement which is contingent upon rule-following which is in concordance with the rule. However, it is interesting to note that one study found that relieving social anxiety in college students with adaptive plies was only effective when the rule was overtly stated by the client and promptly reinforced by the therapist (Zettle

& Hayes, 1979). This would suggest that the effectiveness of plies given in the initial stages of therapy may be greater when *external* listener-mediated reinforcement is the primary controlling variable.

Avoiding the use of tracks in self-instructional approaches, and instead focusing on plies which embody coping strategies and stress inoculation, may increase the potency of SIT (Zettle & Hayes, 1982). That is, it may be the case that using verbal statements to deal with stressful situations operates through verbal encouragement (pliance), as opposed to the accuracy of the description of the natural contingency specified in the rule (track). Such a perspective may represent a practical way to improve SIT approaches (Zettle & Hayes, 1982). This is a relatively straightforward hypothesis. If someone is about to engage in a stressful behaviour, such as placing one's hand in a bucket of ice-cold water, as in stress inoculation for pain, it makes sense that a rule such as "If I put my hand in the water it will feel cold but I can handle it" (an explicit track), would possibly not control most people's behaviour as strongly as "My therapist will be pleased with me and reassure me of my progress if I can endure this discomfort for 15 seconds" (an explicit ply). Put simply, Zettle and Hayes are implying that if self-instructional approaches are designed to maintain responding in the face of aversive consequences, then pliance, which is less affected by these consequences, should control behaviour more reliably.

Accordingly, in an RGB C-B approach, conditions which maximise the likelihood of pliance would be arranged during a self-instructional exercise (e.g., making it apparent to the client that the therapist will monitor and consequate a correspondence between self-instruction and behaviour), yet the rationale for attending to these conditions would be unclear within the cognitive framework of SIT. Furthermore, giving plies to a client is different to training *rule-formulation*, where the client can observe the aspects of a situation and construct a rule accordingly. This would suggest that self-instructional methods such as those used in SIT are most likely to be helpful in treating relatively specific behavioural problems for which certain actions have been identified as effective (Zettle & Hayes, 1982). If one recalls the earlier review of the empirical status of SIT, one can see that this type of therapy *is* most suited to such problems (e.g., test anxiety, pain management).

Finally, another way a therapist might train a client to recite pre-formulated rules would be to give a chain of instructions, perhaps for situations in which the client was needing to deal with more difficult problems. That is, in contrast to giving one rule to be followed, a more complex contingency might require a sequence of rules. This seems to correspond most closely to the SIT approach of problem-solving training (PST) outlined in Part Two. Ironically, although PST has been taken by some to be a "cognitive" approach (e.g., Meichenbaum, 1985), Skinner (1969) frequently described problem-solving as being a product of precurrent verbal behaviour. Nevertheless, the stages that comprise Meichenbaum's (1986) synthesis of problem-solving approaches might be defined as involving the following rules: Step one, "clearly define the problem", can be thought of as describing the most apparent features of the contingency for which there is no immediately available response other than the process of problem-solving. Step two, "create a list of realistically attainable outcomes and for each possible outcome detail the corresponding course of action that it requires", can be described as matching all of the possible consequences with their corresponding course of action. Step three, "compare these plans with some solutions you might expect other people to try in the same situation, and add these to your own list of alternatives if necessary", can be thought of as comparing the self-rules with other previously tried rules and removing them if they conflict, but adding them if they are recognised as better (thus ensuring that the final rule is not a *naïve* rule).

Next, steps four and five, "list the benefits and disadvantages for each alternative, and proceed to rank them in order of preference on this basis", and "rehearse the alternatives through imagery or behavioural practice", can be described as selecting the rule most appropriate to the problematic contingency and privately rehearsing an approximation of compliance (overtly or covertly). Step six, "try out the alternatives that seem most beneficial and most likely to achieve the desired outcome", can be simply described as following the chosen rule. Finally, steps seven and eight "expect failures but reward oneself for trying and repeat the procedure if the tried solution was found to be ineffective" and "reframe the original situation in light of the problem-solving process for future reference" can be thought of as reinforcing oneself for the behaviour of problem-solving in general, even if compliance with the rule in this instance resulted in

punishment or outcomes less rewarding than those specified by the rule. Thus, PST can be adequately described as a *chaining* process, whereby the behaviour of following each rule in the problem-solving sequence acts as discriminative stimulus for the presentation of the next rule.

I have outlined several ways RGB might be incorporated into behaviour therapy, and how these methods might be a way to incorporate several techniques of cognitive therapy in a theoretically consistent manner. The behaviour therapist may well say "that's fine, but what are the implications for my usual routine of behaviour therapy?" As such, we now turn to a discussion of the way therapeutic techniques of RGB might be synthesised with standard contingency management.

3.9 Complementarity Between RGB and Other Behavioural Techniques.

A major advantage in using the concept of RGB to explain the practice and theory of CBT is that it may put researchers in a position to somewhat refine the process of change that it is clearly capable of. That is, the effective ingredients may be more easily identified if those ingredients share a common terminology. Take the treatment of a client whose primary problem was racist behaviour. A CBT program might focus on re-evaluating the core belief that a particular race of people were unintelligent, while another part of treatment might focus on directing the client towards socialising with members of the discriminated race (i.e., a social-interaction contingency). From an integrative angle, a two-tiered approach such as this can be challenged. Are both strategies necessary? How will one complement the other? Could just one of the two strategies be used, so as to free extra therapy time for other issues? If only the social interaction component were needed, could rules be used to facilitate this type of training? If reinterpreting the cognitive components of CBT as RGB is meant to be a solution to this problem, then we need to ask the question "how will these new RGB concepts blend with the techniques and procedures already used in behaviour therapy?"

More specifically, if we are claiming that the concept of RGB can turn CBT into a more parsimonious and internally coherent paradigm for both scientist and practitioner,

then we need to *formalise* how RGB techniques might be used in conjunction with traditional methods of both cognitive and behaviour therapy, in a way that is both logical and complementary. I will now try to clarify exactly how rule-control might aid normal contingency-management. Thus, this section is not related to cognitive therapy per se, and instead concentrates on the behavioural aspects of CBT. Nevertheless, each section is directly related to the three main mechanisms of cognitive change I described earlier, because each may complement the non-verbal components of behaviour therapy in different ways. Specifically, I will look at the following three areas of complementarity: (a) how one can use the elimination of poor rules to boost the learning of new behaviours, (b) how the training of rule-formulation can assist non-verbal functioning, and (c) how the direct presentation of new rules to a client can help non-verbal contingencies control and maintain desirable behaviours.

1. Using the dismissal of poor rules to enhance the learning of non-verbal behaviours. I have discussed ways in which poor rules can be eliminated from a person's environment by directly punishing the presentation of those rules. This may help other aspects of behaviour therapy because other non-verbal behaviours may become available when rules that previously evoked incompatible behaviours have been dropped. For example, if someone abandons the dysfunctional rule "I need to stay at home because something bad will happen to me in an open space", then other needed social behaviours, which require leaving home, can be taught. In articulating rule-abandonment with behaviour therapy, a clinician needs to ask "what rules could be preventing the contingency-shaped behaviours I wish to train, and would it be beneficial to first focus on eliminating the presentation of any hindering rules?" Thus, poor rules can be first seen as a hindrance to normal contingency management.

2. Teaching rule-formulation to assist non-verbal training: I have already discussed several aspects of teaching a client how to formulate adaptive rules, but now I will elaborate on how this might coalesce with traditional behaviour therapy. I first discussed track-formulation, and the ways that accurate tracks could be constructed and tested. However, a client's track-production methods might not be so faulty as is their inability

to choose the *right* contingencies to tact. Thus, traditional behaviour therapy may assist a client to seek out new contingencies, that, once pointed to with a rule, would enable responding to a more productive contingency. For example, a woman in an abusive relationship may be constructing and following tracks which tact relationships between certain behaviours and avoiding battery. However, she may need to tact *new* contingencies surrounding protection and self-preservation, such as leaving home or seeking support. However, this may not be possible unless reinforcement and stimulation for certain behaviours is made available, such as the therapist's making her aware of alternative shelters, or arranging contact with support groups for women. Thus, contingency management may assist the teaching of positive tracking.

Conversely, training in pliance formulation may provide the client with the extra impetus needed to engage in difficult non-verbal behaviours. For example, training the client to construct rules when insensitivity is required may be useful. If a battered woman could independently create such rules as "If I get out of this relationship, I can save my children", then insensitivity to the immediate negative consequences of leaving her husband might be produced, such as those contingencies related to a potentially worsened financial position. Someone may also obtain the required sensitivity by using rules which more clearly tact the natural contingencies pointed to by the rule, such as "If I don't get out of this relationship, my children or I may eventually be beaten to death". Thus, one aspect of training rule-formulation could involve teaching a client to independently use pliance when homework assignments (e.g., contacting a refuge shelter) seem aversive or overly demanding. In summary, in deciding how training rule-formulation can complement behaviour therapy, a therapist needs to ask "would skills in rule-formulation be enhanced by contact with certain contingencies, and can this client independently formulate rules which will provide enough insensitivity to override the immediate aversive consequences of these homework assignments?"

3. Presenting rules to a client in order to assist the training of non-verbal behaviours: If a therapist is willing to offer rules to a client outright, then using rules to induce behaviours which will later come under the control of naturally occurring contingencies may be useful. Using rule-control as a *temporary* method of behaviour change could be

aimed at behaviours for which obtaining long-term maintenance is difficult. That is, direct instructions can be used to "kick-start" contingency-shaped behaviours into action (Baum, 1994). For example, a therapist may direct clients with alcohol problems to say to themselves "Even if I go to this party without being drunk, I can socialise by relaxing and focusing on the people I talk to". A person suffering from alcoholism who recites such a rule before social situations may overcome the initial anxiety caused by sobriety, and soon discover the unspoken and rewarding contingencies of sober intellectual conversation (i.e., contingency-shaped behaviours). Thus, rules may aid a client's engagement with certain contingencies to which they would not otherwise be exposed.

However, one caveat is necessary here. Using rules as a way to direct people into new situations should not obscure the need to *teach* behaviours which could help the client deal with the situation of concern. In fact, a criticism of cognitive therapies is that by failing to view beliefs as rules they pay considerably less attention to ensuring that the client's verbal repertoire (or range of available RGB's) is adequate (Poppen, 1989). The problem with this is that while new beliefs may seem more "rational" or more "logical", they may be directing a client into situations for which she is not adequately prepared to cope (i.e., where she is unfamiliar with the required behaviours). However, an RGB C-B approach would hold that non-verbal techniques and RGB can each work to maximise the utility of the other. Under this philosophy, a client would never be instructed to use a rule unless it had first been established that the ability to execute accurate compliance was adequately learned.

3.10 Empirical problems with the Concept of RGB in an Integrative Context.

Any theoretical analysis is at best speculative without empirical support. However, I have tried to avoid presenting propositions that could not be empirically tested in an applied research setting. Such research would encourage other psychologists to see the merits of adopting the approach, and would certainly serve to encourage the formulation of new and refined hypotheses about RGB in the context of CBT. Unfortunately, there has been a great amount of critical attention drawn to the methodology used in the experimental examination of instructional control, and it may be the case that these issues

need to be resolved before projects similar to this one can be taken seriously. As I will discuss now, the problems with RGB and psychotherapy integration can in part be considered to be empirical.

Empiricism is the philosophical position which argues that truth is exclusively established through sense perception and experience (Gale, 1979). Refined for the purposes of science, empiricists advocate for the *positivist* approach of establishing knowledge through observation and experiment. In this project I have tried to draw attention to some of the experimental findings on RGB, and how these might be used to strengthen the argument for defining cognitive therapy in radical behavioural terms. However, there are problems with using laboratory research to make assumptions about the use of RGB in clinical settings. Much of the research may not be applicable to relatively uncontrolled conditions (i.e., clinical settings). In addition, there are methodological problems associated with these studies that may have implications for Skinner's theory of RGB as a whole.

At the centre of the debate on the empirical validity of RGB is the *verbal control* and *epiphenomenon* theories (Plaud & Newberry, 1996). The verbal control theory is basically the premise that RGB is a major causal variable in human activity and is an important factor in the development of a child's transition from animal to adult human responding. In contrast, the epiphenomenon theory of RGB holds that rule formation is merely a *by-product* of a verbal organism's contingency-shaped behaviour. That is, even if it has been shown that people can explain their behaviour by describing rules, epiphenomenon theory suggests that this occurs *after* responding and does not directly affect it. Thus, in application, the epiphenomenon theory argues that even if there is a high correlation between verbal reports of contingencies and patterns of responding to those contingencies, the two are not causally connected, and rules do not directly affect behaviour.

The claims made by the epiphenomenon theory are not unfounded. For example, the Pouthas et al. (1990) study described earlier concluded that language development is responsible for an increased ability to respond efficiently to schedules of reinforcement. As Galizio (1987) has pointed out, a child learns to respond to a host of new contingencies in the early years of life, often involving complex motor performances and

highly variable schedules (e.g., games, social interaction, toilet training, etc.). Thus, to pick out rule-control as the factor responsible for these changes may be a premature assumption to conclude on the basis of current developmental research.

The rule *report-method* of studying RGB has also been criticised. Humans often inaccurately describe the contingencies they have responded to, and although some studies have found a strong correspondence between verbal report and responding (e.g., Bentall, Lowe, & Beasty, 1985), others have found very little (e.g., Matthews, Catania, & Shimoff, 1985). Furthermore, self-reported rules may simply be retrospective reports of a previously tacted contingency, and although accurate, may not be responsible for verbally directing behaviour (Galizio, 1987). That is, even when a description of a contingency is technically accurate, this does not necessarily mean that continuous presentation of a covert rule is responsible for a repetitive behavioural pattern. For example, I may be able to accurately describe the contingencies that I had to respond to in order to make my breakfast this morning, but this is by no means evidence that my "breakfast making" behaviours were rule-governed. It is a logical error to assume that any self-described regularity in behaviour qualifies as "rule-governed". However, some rule-report studies may use such data as adequate criteria for inferring rule use.

Similar problems with the rule-presentation method have been identified. For example, Galizio (1987) suggests that researchers may mistakenly infer rule use because of an apparent rate of responding which is over and above the normal level of proficiency that would be demonstrated if direct contact with the contingencies had shaped responding. However, previous exposure to similar but hidden contingencies, rather than private rule-recital, may actually be responsible for such responding. According to Galizio (1987), many of these limitations may stem from the fact that research on RGB has often relied on correlational analyses, yet the results are often interpreted as representing causal relationships. This may represent a potentially serious problem, because as a derivative of operant theory the concept of RGB rests on the notion of contiguous causality.

These problems are not easily solved in the context of an assimilative RGB C-B integration. From a methodological viewpoint, it might even be said that inferring the use of rules as the cause of dysfunctional behaviour is as objectionable as attributing those

causes to cognitive structures. Thus, clinical research might first benefit from suggestions as to how researchers can reliably study rule-control, rule formulation, and the study of covert behaviours in general. Some authors (e.g., Taylor & O'Reilly, 1997) have already begun to refine these methods so that may they be less susceptible to the criticisms raised here.

However, it is also clear that if the concept of RGB is to survive as a clinically viable tool it needs to be incorporated into well-controlled process and outcome studies, in addition to basic behavioural research. Of course, all psychotherapists should demand proof that an RGB based therapy actually *worked* before using it. One of the most positive changes in psychotherapy over the last three decades is the increased emphasis on using only therapies which have been shown to be effective under controlled conditions (Bergin & Garfield, 1994). This point is also paramount from an integrative perspective. Even though attempts to integrate psychotherapies have resulted in many new perspectives and innovative clinical theory, the relatively slow progress of the integrative field may be in part attributable to a failure on the behalf of researchers to specify ways that their theories could lead to testable predictions in psychotherapeutic settings (Arkowitz, 1992). According to Arkowitz, providing sound research method that is linked to an empirically verifiable theory is the greatest challenge that psychotherapy integrationists must face if their work is to be taken seriously. Thus, empirical support could be the vital precursor to any such acceptance of an RGB C-B therapy in mainstream clinical psychology.

However, one must not forget that psychotherapy integrationists aim to achieve *different* things depending on their various motivations, and thus the specific aims of each attempt can vary greatly. Nevertheless, one common interest among all integrative attempts is to show that the resulting combination could potentially be better than the individual therapies on which they are based (Arkowitz, 1992). A *constructive* treatment strategy refers to experiments which assess the value of adding new components to a therapy by testing different therapeutic conditions. The use of a constructive design for a comparative outcome study could be useful in comparing an RGB C-B therapy to other modes of treatment, because the method allows support for one of two or more treatments based on conflicting theoretical propositions (Kazdin, 1980). Of course, even if a rule-

based approach was not found to be *universally* more effective than its competitors, outcome research could help to identify the situations in which it was most potent. For example, it may be the case that a rule-based therapy would be particularly effective for individuals who would prefer a treatment which clearly related private events to overt behaviours. This might include those who are unable to accept the premise that underlying "schema" exist "somewhere in their heads", or clients who would simply prefer a treatment which was consistent in philosophy with their own *realist* notion of the world. Ultimately, the value of empirical support for an RGB C-B integration depends on our ability to show, for whatever reason, that such a therapy could be better than purely cognitive, purely behavioural, and standard cognitive-behavioural therapies.

3.11 Theoretical Problems with the Concept of RGB in an Integrative Context.

Rationalism refers to the idea that knowledge can be derived purely through reason and argument (Chaplin, 1985). Because such discovery is said to be somewhat independent of experience and sense perception, rationalism is sometimes set up in opposition to empiricism. The rationalist argues that the basis of any theory should be logically consistent and be grounded on premises which are non-contradictory. Thus, the process of logically defining RGB is closely related to establishing the fundamental premises of the theory. However, some have argued that much of the theoretical work on the concept of verbal behaviour and RGB is plagued with an absence of precise and universal terminology. It also seems possible that these theoretical difficulties could pose as a problem for those wanting to use the concept as an integrative tool.

For example, some have argued that the notion of verbal behaviour itself is poorly defined. Skinner (1957) defined verbal behaviour as "behaviour reinforced through the mediation of others" (p. 14). Some have argued that this definition is somewhat arbitrary when its scope is considered. Consider the following statement given by Hayes (1993):

> If a rat pushes a bar and gets food pellets according to a variable ratio (VR) 10 schedule, this bar pressing is verbal behaviour by Skinners' definition: The reinforcement of bar pressing is socially mediated, and

> experimenters are trained to deliver the food on schedule precisely so as to reinforce the bar pressing of the rat. Skinner recognised this, and suggested that the experimental animal and the experimenter form a "small but genuine verbal community" (Skinner, 1957, p. 108). If even such a simple operant behaviour is "verbal" we have a functional definition that is only relevant at the formal modes of reinforcement delivery (p. 288).

Thus, if the term *verbal-behaviour* is applicable to the simple behaviour of animals, then it may not be the most appropriate definition for its intended use of understanding the behaviour of human interaction. In all fairness, however, it must be pointed out that Skinner probably used the term *verbal* to refer to a *general category of behaviours* rather than as a precise definition that was meant to stand up to a detailed scrutiny of conceptual precision. As Reese (1991) points out, it was only a matter of time before the concept of verbal behaviour was abandoned as a technical term and simply used to refer to a broad class of behaviours:

> I will not give a technical definition of verbal behaviour, mainly because it is not a technical term. "Verbal behaviour" is useful as the name of a book, but Skinner's (1957) book on verbal behaviour and criticisms of it demonstrate that it does not need a precise definition. It is a loosely defined generic label, but even if it were a precisely defined generic label it would still refer to an abstraction. No one can show me a verbal behaviour, but almost anyone can show me an instance of this rough class of behaviours. The problem of psychological interest is not to explain how the class or class name develops, but how instances of the class develop (p. 152).

Perhaps, then, a general term is harmless if it is useful in referring to a particular category of behaviours. The term *eating behaviour* is useful when referring to the functional activity of self-nourishment, yet there are undoubtedly behaviours which are not clearly either "eating" or "not-eating" (e.g., chewing gum). However, if a *defining* characteristic of RGB is that it is verbal behaviour, or in some way involves behaving verbally, then we can see how an absence of such clarity might have contributed to

definitional complaints of the concept of RGB. Baum (1994) also emphasises the importance of establishing a difference in the object of discrimination as a means of defining RGB, as in the following quote: "Rule-governed behaviour depends on the verbal behaviour of another person, whereas contingency shaped behaviour requires no other person, only interaction with contingencies" (p. 131).

But others have expressed the concern that the use of such a distinction between animal and object to define RGB is illogical, because from a strict radical behavioural perspective, there is no reason why animals and humans should produce functionally different discriminative stimuli. Take the following example given by Hayes and Hayes (1989):

> Suppose a dog responds appropriately to its masters request, "go get my slippers". Now imagine the same dog getting the slippers in response to the sound of its master's car in the driveway. It would be bizarre to call the former "rule-governed behaviour" and the latter "contingency-shaped behaviour" based on the means by which the stimulus objects occasioning the behaviour were produced (Hayes & Hayes, 1989, p. 126).

Thus, we can not distinguish RGB from contingency-shaped behaviour simply because it is under the control of "verbal" stimuli, and using the origin of the stimulus as a definition does not seem entirely logical. However, Skinner (1969) did not define RGB as simply "verbal", but as behaviour under the control of contingency-specifying stimuli. This description is sometimes seen as problematic because it appears to pre-assign formal features to a class of stimuli. Such pre-assignment has the potential to be misleading because, as stressed on many occasions in the current essay, rules may not always structurally possess the three elements of a contingency, yet the behaviours they evoke are still what we would call *rule-governed.* As such, in this project I have often referred to rules as only *implicitly* specifying a contingency (i.e., implicit by virtue of *function*).

Of course, radical behaviourism strongly advocates for the use of functional definitions - a rule is a rule if it serves the purpose of a rule. That is, a rule has *specified* a contingency, regardless of its explicitness, if it has induced RGB, *ipso facto*. But this

does not allow the precise identification or differentiation of rule-governance from other "non rule-governed" behaviours. Perhaps a more general definition of RGB as a class of behaviours describable as "compliance with instruction" would be more appropriate. However, an explanation for the term *instruction* still remains wanting; even the simple behaviour of a hen pecking a button may be seen as compliance, but the red light inducing such behaviour would not ordinarily be termed a rule, even if reinforcement for this activity was verbally-mediated by an experimenter. Thus, in contradiction with the general use of the term *rule*, it would seem that even a red light can somehow "specify" a contingency or "instruct" one how to respond to it.

The most likely reason for the absence of a clear distinction between RGB and other contingency-shaped behaviours is that the two exist on a mutual continuum. Rachlin (1992) has pointed this out when reviewing debates on the distinction:

> ....the papers [on RGB] oversimplify things by taking continua and making dichotomies out of them: probability and delay, are both obviously continua (a repeated probability being only a variable rather than fixed delay). But rule-governed and contingency-shaped behaviours also lie on a temporal continuum. All behaviour is in a sense rule-governed, and all behaviour, in a sense, contingency-governed. Even a rat's single reinforced lever presses can be seen as an instance of a very complex set of rules governing muscular movement. Ultimately, even building a house can be seen as a very long-duration operant (Rachlin, 1992, p. 86).

Thus, there appears to be no veridical means of establishing a point along this continuum for a definitional distinction. But one would expect that rule-governance at *least* represents a pattern of behaviour sufficiently distinct from other types of behaviour to warrant a separate label. The arguments documented here would suggest that there is *no* conceptual way to illustrate the uniqueness of RGB.

The point here is that not everyone agrees on what is meant by the term *rule-governed behaviour*, and such confusion breaks from the behavioural tradition of using nothing other than precise statements and empirical definitions. Speaking about the definitional qualities of RGB, Burns and Staats (1991) note that "the concept may be

likened to the social learning theory concept of 'modelling' which is used to label a number of phenomena....When one concept is stretched to consider phenomena with different characteristics, the concept becomes mushy and indefinite" (p. 130). While some have argued that rule-governed behaviour is hard to define because there is "no set of common features that may define a functionally distinct category" (Cerutti, 1989, p. 261), others continue to maintain that the definitional problems stem from an inability to determine exactly how rule-governed behaviours differ from other contingency-shaped behaviours (c.f. Schlinger, 1993). Some are even doubtful as to whether there is an adequate basis for a tenable theory. As Brownstein and Shull (1985) note, "Unless the interpretations of complex human phenomena derive from a careful, detailed and rigorous use of technical terms and concepts, it will be hard to learn whether or not the interpretations [of RGB] are ever satisfactory" (p. 267).

As it relates to this thesis, such criticism of RGB theory begs the question: "what are the ramifications of using an undefined concept as a central theme for a theoretical integration of two different paradigms of therapy?" First, it could be said that the greatest difficulty lies in the fact that any therapy based on a set of guiding assumptions which do not rationally cohere is likely to be disorganised and eventually replaced by theory which produces fewer conceptual anomalies. This is a serious consideration because it is unlikely that practitioners trying to grasp the theoretical basis of an RGB C-B clinical technique would place confidence in an analysis which had fundamental problems with its most elementary concept. Another problem lies in the hindering of theoretical advancements at the academic level. It can be shown that much of the theoretical literature on RGB resolves around problems of inconsistency in the definition of verbal behaviour and rule-governance (Burns & Staats, 1991). However, these discussions are in the hands of a relatively small group of radical behaviourists. History has shown us that revolution in psychology is a slow and laborious process that requires the work of many, and an assimilative RGB C-B integration is essentially aiming at such a revolution. It seems highly unlikely to me that any such academic collaboration would seem inviting when the hoary issue of defining verbal behaviour and rule-governance would necessarily precede any fruitful theoretical advancements.

Tangentially, it is probably fair to say that from an *empirical* angle, the absence of a universal definition is unlikely to be problematic. After all, the problem seems to be primarily one of parameters, in that no proposed definition of RGB has clear enough boundaries to prevent the confusion that arises when verbal and non-verbal behaviours seem indistinguishable. Because the definition of RGB used in experiments with instructional variables lies well within those parameters, the absence of a clear definition is unlikely to affect laboratory experiments or the generalisation of basic human research on RGB to clinical settings (i.e., in any experiment which uses instruction as an independent variable, the rule is clearly a *rule,* and definitional problems are too abstract to cause difficulty in controlled laboratory situations). Nevertheless, it is perhaps unreasonable for radical behaviourists to expect cognitive psychologists to even consider a rule-governed interpretation of cognitive therapy when there is not even internal agreement as to what rule-governed behaviour *actually is*. If these definitional issues are not resolved, an immersion of RGB theory into the field of cognitive therapy might be passed off by some as nothing but a contamination of the otherwise tidy area of cognitive clinical science.

In short, the dilemma stems from the suspicion that deficiencies in a conceptual tool will indirectly surface as conceptual difficulties in the finished product for which that tool has been employed. For example, it could be argued that the *gap-in-time* dilemma described earlier has not been fully resolved, and that this might limit the use of a radical behavioural approach to cognitive therapy. This of course would not have implications for psychotherapy - we certainly know that rules affect behaviour and can be used to implement behaviour change, regardless of any "gap-in-time". However, in the context of giving a *radical behavioural account of cognitive psychotherapy with the concept of RGB*, we may face a conceptual barrier. An inferred cognitive structure such as the "acquisition of a schematic rule", forever present and at the ready, does not fall prey to such deterministic problems as temporally distant effect. Yet a radical behavioural analysis would not tolerate the use of a mentalist concept to explain away the problem of gaps-in-time, and an appeal to memory, core schema, beliefs, or any other hypothetical construct is simply not an option. To ignore this issue would run the risk of letting a

radical behavioural concept become essentially cognitive, which would undoubtedly undermine the philosophy of an assimilative integration.

The gap-in-time dilemma has not posed as a problem to this analysis. Here I have viewed rules as being *only* those that act as immediate verbal antecedents to RGB. This has necessarily meant that if rules constructed in therapy are expected to control behaviours outside of the therapeutic session at a later date, then they may do so only as precurrent behaviours which have occurred as immediate verbal antecedents. In accordance with this, here I have assumed that if private and self-rule control is to replace cognitive structures, then there should be an emphasis on training this precurrent behaviour, rather than simply presuming that it occurs without some external aid. Thus, there has been an emphasis on (a) training the client to recognise situations where a rule should be stated, (b) maximising the conditions that would allow the private-recital of rules to generalise to new situations (e.g., recreating conditions in therapy which approximate the situation for which the private rule-recital is intended), and (c) ensuring that there is that provision of reinforcement for RGB in *connection* with the recitation of the rule (e.g., having a client say to herself "see, once I said that rule I knew I could do it"). Working around the gap-in-time issue in this way may seem like an overreaction to what may seem like an overly-theoretical and practically trivial matter. But the only alternative is to *presume* that rules can have non-contiguous effects, because Skinnerian theory simply does not explain how this might occur within the rules of our science.

However, these controversial issues are quickly disappearing from contemporary behavioural literature, as behaviour analysts continue to develop and explore uncharted areas of human psychology. Currently, some are adopting an entirely new conceptualisation of RGB, what Hayes (1994) has called *Relational Frame Theory* (RFT). RFT has ultimately been derived from early research on stimulus equivalence (Sidman & Tailby, 1982). Stimulus equivalence research involves training participants to match sample stimuli to comparison stimuli based on derived relationships. Sidman and Tailby originally identified three types of sample-comparison stimulus relationships that could be learned: *reflexivity, symmetry, and transitivity*. Reflexivity involves matching *identical* stimuli. For example, "given A1, select A2"; "given B1, select B2", and so on. Symmetry involves learning a *two-way* relationship between stimuli. For example, if a

participant has learned "if given A, then select B" and then comes to "select A, when given B", then symmetry has been shown. Transivity involves learning a sample-comparison relationship derived through an association with a *common* stimulus. For example, after being trained to "select B, given A", and "select C, given B", transitivity has been shown when the subject has come to "select C, given A". According to Sidman and Tailby, complete stimulus equivalence is shown when all three relations have been learned. Hayes has proposed that examples of stimulus equivalence are a defining property of verbal behaviour, and RFT involves a broader account of stimulus equivalence.

According to Hayes (1994), learning arbitrarily applicable relations between stimuli is itself a class of operant behaviour, and can involve learning generalised stimulus sets for which an infinite number of items can be included or "framed". For example, the arbitrarily applicable relationship "bigger than" can be learned as a relational frame for which an infinite number of size-related stimuli can be included. That is, in any context that selects the behaviour of "pick the biggest", any presenting differentially-sized objects could be included as being part of this relational frame. Hayes (1994) has suggested that there are three major types of relational responding: *mutual entailment*, *combinatorial entailment*, and *transformation of stimulus function.* Mutual entailment involves a reciprocally-dependent relationship between stimuli. That is, if A is in some way related to B, then B, by derivation, is conversely related to A. For example, if the apple pie is hotter than the blueberry pie, then in a context which selects "the cooler of", one can select the blueberry pie through a derived relation. Combinatorial entailment involves a relation derived through a common stimulus. That is, in a context in which A is associated with B, and B is associated with C, a relationship derived between A and C is learned. For example, if ice is water, and water is H2O, ice can be selected as H2O through a derived relation between their mutual relationship to water. Transformation of stimulus function occurs when an additional psychological function of a stimulus has been derived in consistency with its mutual relation to another. That is, in a given context, if A is mutually related to B, and A acquires an additional discriminative property, B will acquire that property consistent with its relation to A. According to Hayes, RGB is essentially the behaviour of "framing-relationally": a contextually-

controlled, arbitrarily-transferable pattern of responding to stimuli through a derived relation.

RFT is not entirely incompatible with Skinner's notions of rule governance (Hayes & Wilson, 1993). A rule in the present sense would most likely be construed as a complex relational frame of the type "in the occasion of X, do Y, and Z will occur" (i.e., any number of natural events could be substituted for X, Y, or Z). Hayes (1994) has also suggested how the concepts of pliance and tracking can be incorporated into RFT. More importantly, RFT may in some aspects be less vulnerable to the definitional criticisms to which Skinner's discrimination theory of RGB has been subjected. For example, the distinction between verbal and non-verbal organisms is not an issue with RFT, because it has been simply shown that animals cannot respond to arbitrarily applicable relations, and thus the animal-human distinction is not a concern (Hayes & Wilson, 1993). Furthermore, by relying on the notion of higher-order relational discriminations, as in RFT, the gap-in-time dilemma can be avoided by defining a rule as a stimulus which has demonstrated a *function-altering capacity* (Schlinger, 1993). That is, rules themselves are not directly discriminated, rather they have the ability to *alter* or *establish* the stimulus control of contexts or objects occurring in another time or place. Consider the following rule; "stop at a red light and you will avoid crashing". Schlinger would argue that if someone were to stop at a red light after being told this rule on an earlier occasion, then the rule has altered the discriminative function of the red light because the rule evoked a behaviour that could not have occurred without the prior presentation of that rule event (i.e., as opposed to the rule itself functioning as a stimulus which is temporally distant from its evoked behaviour). Subsequently, Schlinger argues that if the term *rule* is to be used it all, then it should be reserved for events that act as *function-altering stimuli* (FAS).

Many new behavioural theories may eventually prove to be effective in assisting the assimilation of cognitive techniques into a behavioural analysis. Although there have been advancements in the development of a general mode of psychotherapy based on RFT (see Hayes, 1989), the concept has not yet been employed as a way to theoretically integrate the cognitive and behavioural therapies. Nevertheless, future research may see an employment of RFT theory for this use. Of course, viewing rules as relational frames

or as FAS are only two of many possible ways to deal with the concept of RGB and its conceptual problems, and can at best be considered incomplete theories at this point in time. The pivotal issue here is that notions of rule-governed behaviour, internal dialogue, thoughts and the like are being understood within a parsimonious (albeit complex) framework which does not require the huge jump to mentalism that other schools have made, using nothing other than the core tenets of operant theory as guiding themes.

3.12 Final Conclusion

In this project I have tried to attend to three major areas: (a) developments in the field of psychotherapy integration that can assist in specifying both what this type of cognitive-behavioural rapprochement should aim to achieve and what possible pitfalls should be avoided, (b) empirical strengths of the concept of RGB which can serve to validate the argument for an assimilative RBG C-B integration, and (c) the need to evaluate such an integration as an emergent psychotherapeutic system, so as to ensure that any developments in an RGB C-B integration are both thorough and clinically applicable. However, being comprehensive does not necessarily ensure ultimate success, and there are problems associated with psychotherapy integration which lie outside of those caused solely by problems with the concept of RGB. I will now present some of these difficulties before concluding Part Three with a final note about their relation to the broader aims of this thesis.

Unfortunately, there is the potential to cause damage when integrating two therapies. Probably the greatest concern is the risk of obscuring valuable factors unique to a therapy when adjusted for the purpose of integration (O'Donohue & McKelvie, 1993). In the context of assimilative integration, this would imply that when a therapy is immersed into a foreign theoretical context, it should retain *all* of its former value (presumably if something *must* be lost on immersion, it should be justified by the promise of greater overall gain). The model proposed here is by no means excused from this potential criticism. Although many elements of cognitive therapy were discussed (e.g., the applicability of Beck's (1976) distortions in information-processing to problems in rule

formulation, a rule-governed interpretation of the Socratic method of questioning), a huge amount of clinical cognitive literature was omitted. Furthermore, although space limitations forced an omission of many unique aspects of REBT, CT, SIT, and social-cognitive theory, it is unclear as to whether some components of cognitive therapy are even *amenable* to an analysis in terms of RGB. For example, how would the useful distinction between *schema* and *beliefs* be interpreted as RGB? Could the concept of RGB be used to explain automatic thoughts, cognitive shifts, repressed memory, and so on? Hundreds of other examples could be provided, and even if many of these components could be adequately described using an operant framework, it is not always clear why these reinterpretations would be expected to be more useful than their cognitive prototypes (e.g., it could be argued that my behaviour-analytic translation of Meichenbaum's (1985) synthesis of PST is *harder* to understand than his). As Laudan (1977) rightly points out, "because the assessment of competing theories is a multi-factorial affair, parity with respect to one factor in no way precludes a rational choice based on disparities at other levels" (p. 166). Put simply, the fact that *some* aspects of cognitive therapy *can* be defined as RGB does not mean that an RGB C-B integration is *totally* justified.

However, this raises an important point in regards to the intended purpose of the integrative account offered here, and also supports my decision to choose the assimilative approach over other integrative methods. This thesis never claimed that cognitive therapies should be seen as *fully* explainable in terms of RGB, nor was I proposing that any cognitive concept that can be defined as RGB must be considered as *improved.* Rather, the guiding assumption underlying this project has been the contention that by using the radical behavioural concept of RGB, behaviourists may move toward a logical and justifiable incorporation of *some* of the merits of cognitive therapy which may have been previously overlooked. Thus, the interpretations of cognitive lore as RGB given here should not be considered as finished products, nor should they be considered as exhaustive - it may be the case that *other* radical behavioural concepts can also assist in redefining cognitive therapy, and in some cases, be utilised where RGB theory has fallen short. To leave this point unnoticed would risk appearing as though the rationale for my assimilative integration was primarily born out of a defence of the radical behavioural

school, or what Zettle (1990) has referred to as meeting the "cognitive challenge". However, such defence is secondary to the potential gains of actually *employing* the radical behavioural approach in the development of new psychotherapeutic theory.

In Part One I stated that a major driving force behind an assimilative RGB C-B integration was the hypothesis that any paradigm free from conflicting fundamentals is likely to develop more rapidly. How might this be applied to the current project? As I progressed through the basic areas essential to a rule-based psychotherapy (i.e., first defining the working concepts, then outlining general units of pathological speaker and listener behaviour, followed by an introduction of various mechanisms of change), I drew from the necessary cognitive theory needed, or seemingly most suitable. As such, one can see how the *boundaries between cognitive therapies themselves* might slowly disappear when seen in terms of verbal behaviour, because different areas of the rule-based approach are catered for by differing cognitive therapies. This is not an unwanted side-effect. For example, why would there be a need to choose between CT and social-cognitive theory if one could use an alternative approach which incorporated the effective ingredients of both? Integrating the cognitive therapies with RGB may be the easiest way to foster the unification needed for a *truly* integrated CBT.

In fact, at a more general level the issue of parsimony can not be understated when reviewing the reasons for using a radical behavioural approach to psychotherapy integration. In contrast to popular misconceptions, radical behaviourists *are* different to methodological behaviourists, and *can* deal with the notions of thoughts, feelings, and emotions using the minimal repertoire of terms available to them (e.g., reinforcement, punishment, stimulus control). In fact, if one randomly selects any text based upon radical behavioural lore, the terms and concepts are generally based around the same basic tenets underlying operant theory relied on here. One is most unlikely to be guaranteed such reliability when casually browsing through a cognitive treatise, the florid models of which can take almost any shape or form. In choosing a method to integrate two therapies which are at risk of becoming excessively diverse in practice, parsimony is surely an important consideration.

Of course, another way to foster this type of parsimony is to be clear about the *role* of any added techniques in the context of the new parent therapy. Although I made some

initial speculation on the ways RGB and contingency management might work to complement each other, many more questions remain to be answered. For example, could rules be used to assist desensitisation and other techniques based on classical conditioning? To what extent do non-verbal behaviours play a role in rule-construction? Can motor rules (i.e., rules based on observing one's performance) exert as much stimulus control over behaviour as rules based on other verbal behaviour, or verbal origin rules? It is likely that research which examined the extent to which various assimilated factors increased positive outcome in therapy would provide this type of insight. More specifically, deconstructive outcome research may more precisely specify how traditional contingency management could benefit from looking to cognitive therapy and its rule-based alternative for new ideas about behaviour change.

I have offered a brief analysis of how both cognitive and behavioural therapy might be placed within a common theoretical framework. Although it is both speculative and generalised, the motives behind this work are clear and based on logic, reasoning, and empirical evidence. By separating the substantive from the substantial differences between cognitive and behavioural therapy, researchers will be in a position to identify what is really "at work" in CBT. Some would suggest that the cognitive and behavioural therapies may simply be inducing the same psychological processes, but through differing conceptualisations of human behaviour and thought (Arkowitz, 1992). That is, it may not be so much a case of "who's right?", as much as it is a case of "what could be the core reasons that make cognitivists and behaviourists *both* right?" One plausible answer to this question is the possibility that ontological and epistemological differences are needlessly separating two therapies which in effect have a great deal in common. If this is the case, then a theoretical review of the utility of the concept of RGB as an integrator may be the vital precursor to a true integration of these fields. In the same vein, the empirically supported effectiveness of the CBT approach in general should not dissuade researchers from trying to understand the theoretical boundaries that lie between the cognitive and behavioural approaches, because such boundaries may in effect be *limiting* the potential effectiveness of CBT. As Laudan (1977) states:

> If we look at the reception of Darwin's evolutionary biology, Freud's psychoanalytic theories, Skinner's behaviourism, or modern quantum mechanics, the same pattern repeats itself. Alongside of the rehearsal of empirical anomalies and solved empirical problems, both critics and proponents of a theory often evoke criteria of theoretical appraisal which have nothing whatever to do with a theory's capacity to solve the empirical problems of the relevant scientific domain. A broader view concerning the nature of problem-solving - one which recognises the existence of conceptual problems - puts us in a position to understand and describe the intellectual interaction that can take place between defenders of theories which are equally supported by the data (Laudan, 1977, p. 112).

It is in the spirit of Laudan's (1977) emphasis on conceptual review as being central to paradigmatic commensurability that one should remain receptive to theoretical progress with the concept of RGB, as it is these advances which might serve to engender the authority that radical behaviourism could one day lay claim to in the area of semantic and cognitive psychology.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington DC: Author.

Arkowitz, H. (1989). The role of theory in psychotherapy integration. *Journal of Integrative and Eclectic Psychotherapy, 8,* 8-16.

Arkowitz, H. (1992). Integrative theories of psychotherapy. In: D.K Freedheim (Ed.), *History of Psychotherapy: A century of change* (pp. 261-303).WA: American Psychiatric Association.

Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behaviour change. *Psychological Bulletin, 84,* 191-215.

Bandura, A. (1977b). *Social learning theory.* NJ: Prentice Hall.

Bandura, A. (1983). Self-efficacy determinants of anticipated fears and calamities. *Journal of Personality and Social Psychology, 45,* 464-469.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* NJ: Prentice Hall.

Baum, W. (1992) For parsimony's sake: Comments on Malott's "A theory of rule-governed behaviour and organisational management". *Journal of Organizational Behaviour Management, 12,* 2.

Baum, W. (1994). *Understanding Behaviourism: Science, behaviour, and culture.* New York: Harper Collins.

Baum, W. (1995). Rules, culture and fitness. *The Behaviour Analyst, 18,* 1-21.

Beck A.T (1967). *Depression: Clinical, experimental and theoretical aspects.* New York: Harper Collins.

Beck, A.T. (1976). *Cognitive therapy and the emotional disorders.* New York: New American Library.

Beck, A.T. (1993). Cognitive therapy: Past, present and future. *Journal of Consulting and Clinical Psychology, 2,* 194-198.

Beck, A.T., Epstein, N. & Harrison, R. (1983). Cognition's, attitudes and personality dimensions in depression. *British Journal of Cognitive Psychotherapy, 1,* 1-16.

Beck, A.T. & Weishaar, M.E. (1995) Cognitive therapy. In: R.J. Corsini & D. Wedding (Eds.). *Current Psychotherapies* (4th ed, pp. 285-320). Illinois: Peacock.

Benjafield, J.G. (1997). *Cognition.* NJ: Prentice Hall.

Bentall, R.P., Lowe, C.F. & Beasty, A. (1985). The role of verbal behaviour in human learning: II. Developmental differences. *Journal of the Experimental Analysis of Behaviour, 43,* 165-181.

Bergin, A.E. & Garfield, S.L. (1994). Overview, trends, and issues. In: A.E. Bergin & S. L. Garfield (Eds.). *Handbook of psychotherapy and behaviour change* (4th ed., pp. 821-831). New York: Wiley.

Brownstein, A.J. & Shull, R.L. (1985). On terms: A rule for the use of the term "rule-governed behaviour". *The Behaviour Analyst, 8,* 265-267.

Burns, G.L. & Staats, A.W. (1991). Rule-governed behaviour: Unifying radical and paradigmatic behaviourism. *The Analysis of Verbal Behaviour. 9,* 127-143.

Buskett, W.F. & Miller, H.L. (1986). Interaction between rules and contingencies in the control of human fixed-interval performance. *The Psychological Record, 36,* 109-116.

Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). Comparison of behaviour therapy and cognitive-behaviour therapy in the treatment of generalised anxiety disorder. *Journal of Consulting and Clinical Psychology, 59,* 167-175.

Catania, C.A., Matthews, B.A. & Shimoff, E. (1982). Instructed versus shaped human verbal behaviour: Interactions with non-verbal responding. *Journal of the Experimental Analysis of Behaviour, 38,* 233-248.

Catania, C.A. (1998). *Learning* (4th ed.) NJ: Prentice Hall.

Cerutti, D.T. (1989). Discrimination theory of rule-governed behavior. *Journal of the experimental analysis of behavior. 51,* 259-276.

Chaplin, J.P. (1985). *Dictionary of Psychology.* New York: Laurel.

Chadwick, P.G.J. & Lowe, C.F. (1990). Measurement and modification of delusional beliefs. *Journal of Consulting and Clinical Psychology, 58,* 225-232.

Clark, D.M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy, 24,* 461-470.

Clark, D.M., Salkovskis, P.M., Hackmann, A., Middleton, H., & Gelder, M (1994). A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759-769.

Clarkson, P. (Ed.), (1997). *Counselling Psychology: Integration of Theory, Research, and Supervised Practice.* New York: Routledge.

Davis, F. W., & Yates, B. T. (1982). Self-efficacy expectancies versus outcome expectancies as determinants of performance deficits and depressive affect. *Cognitive Therapy and Research, 6,* 23-35.

Dember, W. N . (1974). Motivation and the cognitive revolution. *American Psychologist, 29,* 161-168.

Dobson, K.S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 57,* 414-419.

Dollard, J., & Miller, N.E. (1950). Personality and Psychotherapy: An analysis in terms of learning, thinking, and culture. New York: McGraw-Hill.

D'Zurilla, T.J. & Goldfried, M.R. (1971). Problem-solving and behaviour modification. *Journal of Abnormal Psychology, 78.* 107-126.

Ellis, A. (1962). *Reason and emotion in Psychotherapy.* NJ: Stuart.

Ellis, A. (1995a). Changing rational-emotive therapy (RET) to rational-emotive behaviour therapy (REBT). *Journal of Rational-Emotive & Cognitive Behaviour Therapy, 13,* 85-89.

Ellis, A.E. (1995b). Rational-Emotive Behaviour Therapy. In: R.J. Corsini & D. Wedding (Eds.) *Current Psychotherapies* (4th ed, pp. 189-215). Illinois: Peacock.

Emmelkamp, P.M. (1994) Behaviour therapy with adults. In: A.E. Bergin & S. L. Garfield (Eds.), *Handbook of Psychotherapy and Behaviour Change* (pp. 379-428). New York: Wiley.

Feltham, C. (1997). *Which psychotherapy? Leading exponents explain their*

differences. London: Sage.
Frank, J.D. (Ed.). (1961). *Persuasion and Healing*. Baltimore: John Hopkins University Press.
French, T.M (1933). Interrelations between psychoanalysis and the experimental work of Pavlov. *American Journal of Psychiatry*, 89, 1165-1203.
Gale, G. (1979). *Theory of Science. An introduction to the history, logic and philosophy of science*. New York: McGraw-Hill.
Galizio, M. (1979). Contingency-shaped and rule-governed behavior: Instructional control of human loss avoidance. *Journal of the Experimental Analysis of Behaviour, 31,* 53-70.
Galizio, M. (1987). Interpretation vs. experimentation in the experimental analysis of behaviour. *Psychological Record, 37,* 11-15.
Gold, J.R. (1996). *Key concepts in psychotherapy integration*. New York: Plenum Press.
Goldiamond, I. (1976). Self-reinforcement. *Journal of the experimental analysis of behaviour, 9,* 509-517.
Goldfried, M.R. (1983). A behaviour therapist looks at rapprochement. *Journal of Humanistic Psychology, 23,* 97-107.
Hayes, S.C. (1994). Relational frame theory: A functional approach to verbal events. In: S.C. Hayes, J.L. Hayes, M. Sato, & K. Ono. (Eds.), *Behaviour Analysis of Language and Cognition. The 4th International Institute on Verbal Relations* (pp. 9-30). Reno, NV: Context Press.
Hayes, S.C., Kohlenberg, B. S & Melancon, S.M. (1989). Avoiding an altering rule control as a strategy of clinical intervention. In: S. C. Hayes (Ed.), *Rule-governed behaviour: Cognition, contingencies and instructional control* (pp. 359-385). New York: Plenum Press.
Hayes, S.C. & Hayes, L.J. (1989). The verbal action of the listener as a basis for rule-governance. In: S. C. Hayes (Ed.), *Rule-governed behaviour: Cognition, contingencies and instructional control* (pp. 154-190). New York: Plenum Press.
Heimberg, R.G. & Lebowitz, M.R. (1992). *A multi-centre comparison of the efficacy of phenelzine and cognitive-behavioural group treatment for social phobia.* Research presentation at the annual meeting of the anxiety disorders Association of America, Houston.
Hollon, S.D, & Beck, A.T. (1994). Cognitive and cognitive-behavioural therapies. In: A.E. Bergin & S. L. Garfield (Eds.) *Handbook of Psychotherapy and Behaviour Change*. (4th ed., pp. 428-461). New York: Wiley.
Hoyt, F. (Ed.). (1998). *Constructive Therapies.* New York: Guilford Press.
Hussian, R.W. & Lawrence, P.S. (1978) The reduction of test, state, and trait anxiety by test-specific and generalised stress inoculation training. *Cognitive Therapy and Research, 2,* 25-37.
Jaremko, M.E. (1987). Cognitive-behaviour modification: The shaping of rule-governed behaviour. In: W. Dryden, & W. Golden (Eds.), *Cognitive-Behavioural Approaches to Psychotherapy* (pp. 31-60). Cambridge: Hemisphere Publishing Corp.
Jensen J.P, Bergin, A.E & Greaves, D.W. (1990). The meaning of eclecticism: new survey and analysis of components. *Professional Psychology: Research and Practice, 21,* 124-130.

Kanfer, F.H., Karoly, P., & Newman, A. (1975). Reduction of child's fears of the dark by competence related and situational threat-related verbal cues. *Journal of Consulting and Clinical Psychology, 43,* 251-258.

Karasu, T.B. (1986). Specificity versus non-specificity. *American Journal of Psychiatry, 143,* 687-695.

Kazantzis, N., & Deane, F. P. (1998). *Comparing theoretical orientations of New Zealand and North American psychologists.* Manuscript submitted for publication. Massey University, NZ.

Kazdin, E. (1980). *Research design in clinical psychology.* New York: Harper & Row

Kazdin, A. E., Bass, D,. Siegel, T., Thomas, C. (1989). Cognitive-behavioral therapy and relationship therapy in the treatment of children referred for antisocial behavior. *Journal of Consulting & Clinical Psychology, 57,* 522-535.

Keefe, F.J., Caldwell, D.S. Williams, D.A., Gil, K.M., Mitchell, D., Robertson, C., Martinez, S. Nunley, J., Beckham, J.C., Crisson, J.E., & Helms, M. (1990). Pain coping-skills training in the management of osteoarthritic knee pain: A comparative study. *Behaviour Therapy, 21,* 49-62.

Kelly, G. (1955). *The psychology of personal constructs.* New York: Norton.

Kurzweil, R. (1985). What is artificial intelligence anyway? *American Scientist, 73,* 258-264.

Latimer, P.R., & Sweet, A.A. (1984). Cognitive vs. behavioural procedures in cognitive behaviour therapy: A critical review of the evidence. *Journal of Behaviour Therapy and Experimental Psychiatry, 15,* 9-22.

Laudan, L. (1977). *Progress and its problems.* Berkeley: University of California Press.

Lazarus, A.A. (1967). In support of technical eclecticism. *Psychological Reports, 21,* 415-416.

Lazarus, A.A. (1981). *The practice of multi-modal therapy.* New York: McGraw-Hill.

Lazarus, A.A. & Messer (1991). Does chaos prevail ? An exchange on technical eclecticism. *Journal of Psychotherapy Integration, 1,* 143-158

Linton, S.J., Bradley, L.A., Jensen, I. Spangfort, E., & Sundell, L. (1989). The secondary prevention of low back pain: A controlled study with follow-up. *Pain, 36,* 197-207.

Lowe, C.F., Beasty, A., & Bentall, R.P. (1983). The role of verbal behaviour in human learning: Infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behaviour, 39,* 157-164.

Mace, C.F. (1994). Basic research needed for stimulating the development of behavioural technologies. *Journal of the Experimental Analysis of Behaviour, 61,* 529-550.

Mackay, M. & Fanning, P. (1987). *Self-esteem: A proven program, of cognitive Techniques for assessing, improving, and maintaining your self-esteem.* California: New Harbinger.

Mahoney, M.J. (1974). *Cognitive and Behaviour Modification.* Cambridge: Ballinger.

Mahoney, M.J. & Lyddon, W.J. (1988). Recent developments in cognitive approaches to counselling and therapy. *The Counselling Psychologist, 16,* 190-234.

Malott, R. (1988). Rule-governed behaviour and behavioural anthropology. *The Behaviour Analyst, 11,* 181-203.

Manning, M. M., & Wright, T. L. (1983). Self-efficacy expectancies,

outcome expectancies, and the persistence of pain control in childbirth. *Journal of Personality and Social Psychology, 45,* 421-431.

Marlatt, A. A., Baer, J. S., & Quigley, A. A. (1995). Self-efficacy and addictive behaviour. In: A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 289-316). New York: Cambridge University Press.

Marr, M.J. (1984). Conceptual approaches and issues. *Journal of the Experimental Analysis of Behaviour, 42,* 353-362.

Martin G, & Pear, J (1996). *Behaviour modification: What it is and how to do it.* NJ: Prentice Hall.

Matthews, B.A., Catania, A,C., & Shimoff, E .(1985). Effects of uninstructed verbal behaviour on non-verbal responding: Contingency vs. performance descriptions. *Journal of the Experimental Analysis of Behaviour, 43,* 155-164.

Matthews, B.A., Shimoff, E., Catania, A.C. & Sagvolden, T. (1977). Uninstructed human responding: Sensitivity to ratio and interval contingencies. *Journal of the Experimental Analysis of Behaviour, 27,* 453-467.

Meichenbaum, D. H., (1977). *Cognitive-behaviour modification: An integrative approach.* New York: Plenum.

Meichenbaum, D.H., & Goodman, J. (1971) Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology, 77,* 115-126.

Meichenbaum, D.H. (1985). *Stress inoculation training.* New York: Pergamon

Messer, S.B. (1986). Behavioural and psychoanalytic perspectives an therapeutic choice points. *American Psychologists,* 1261-1272

Moe, K. O., & Zeiss, A. M. (1982). Measuring self-efficacy expectations for social skills: A methodological inquiry. *Cognitive Therapy and Research, 6,* 191-205.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton Century Crofts.

Newman, B., Buffington, D.M., & Hemmes, N.S. (1995). The effects of schedules of reinforcement on instruction following. *The Psychological Record, 45,* 463-476.

Novaco, R.W. (1979). The cognitive regulation of anger and stress. In: P.C. Kendall & S.D. Hollon (Eds.), *Cognitive-behavioural interventions: Theory, research and procedures* (pp. 112-169). New York: Academic Press.

O'Donohue, W. & McKelvie, M. (1993). Problems in the case for psychotherapeutic integration. *Journal of Behaviour Therapy and Experimental Psychiatry, 24,* 161-170.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66,* 543-578.

Plaud, J. & Newberry, D. (1996). Rule-governed behaviour and pedophilia. *Sexual abuse: A Journal of Research and Treatment, 8,* 143-158.

Poppen, R.L. (1989). Some clinical implications of rule-governed behaviour. In: S.C. Hayes (Ed.), *Rule-governed behaviour: Cognition, contingencies and instructional control.* (pp. 325-355). New York: Plenum Press.

Pouthas, V., Droit, A. Jacquet, Y., & Wearden, J.H. (1990). Temporal differentiation of response duration in children of different ages: Developmental changes in relations between verbal and non-verbal behaviour. *Journal of the Experimental Analysis of Behaviour. 53,* 21-31.

Rachlin, H. (1992). An important first step, but not the last word on rule-governed behaviour and OBM: Comments on papers by Malott and Malott, Shimamune, and Malott. *Journal of Organizational Behaviour Management, 12,* 85-89.

Reed, S.K. (1992). *Cognition* (4th Ed.). California: Brooks/Cole Publishing.

Reese, H.W. (1989). Rules and rule-governance: Cognitive and behaviouristic views. In: S. C. Hayes (Ed.), *Rule-Governed Behaviour: Cognition, Contingencies and Instructional Control.* (pp. 3-74). New York: Plenum Press.

Resick, P,A., Jordan, C. G., Girelli, S. A., Hutter, C. K., & Marhoefer-Dvorak, S. (1988). A comparative outcome study of behavioral group therapy for sexual assault victims. *Behavior Therapy,* 19, 385-401.

Reynell, J.K. (1977). *The Reynell Developmental Language Scale* (Rev. ed.) London: N.F.E.R.

Roberts, R.N. & Tharp, R.G. (1980). A naturalistic study of children's self-directed speech in academic problem-solving. *Cognitive Research and Therapy, 4,* 341-353.

Sanders, M.R., Rebgetz, M., Morrison, M., Bor. W., Gordon, A., Dadds, M. & Sheperd, R. (1989). Cognitive-behavioural treatment of recurrent non-specific abdominal pain in children: An analysis of generalization, maintenance and side effects. *Journal of Consulting and Clinical Psychology, 57,* 294-300.

Schact, T.E. (1984). The varieties of integrative experience. In: H. Arkowitz and S.B. Messer (Eds.) *Psychoanalytic therapy and behaviour therapy: Is integration possible?* (pp. 107-132). New York: Plenum

Schlinger, H. (1993). Separating discriminative and function altering effects of verbal stimuli. *The Behaviour Analyst, 16,* 9-23.

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

Silverman, M.S. McCarthy, M. & McGovern, T. (1992). A review of outcome studies on rational-emotive therapy from 1982-1989. *Journal of Rational-Emotive and Cognitive Therapy, 10,* 111-175.

Skinner, B.F. (1938). *The behaviour of organisms.* New York: Apple-Century Crofts.

Skinner, B.F. (1945). The operational analysis of psychological terms. *Psychological Review, 52,* 270-277.

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

Skinner, B.F. (1969). *Contingencies of reinforcement: A theoretical analysis.* New York: Appleton.

Skinner, B.F. (1974). *About behaviourism.* New York: Vintage.

Skinner, B.F. (1984). An operant analysis of problem-solving. *The Behavioural and Brain Sciences, 7,* 583-613.

Skinner, B.F. (1986). The evolution of verbal behaviour. *Journal of the Experimental Analysis of Behaviour, 45,* 155-122.

Sweet, A.A. & Loizeaux, A.L. (1991). Behavioural and cognitive treatment methods: A

critical comparative review. *Journal of Behaviour Therapy and Experimental Psychiatry, 22,* 159–185.

Taylor, I. & O'Reilly, M.F. (1997). Toward a functional analysis of private verbal self-regulation. *Journal of Applied Behaviour Analysis, 30,* 43-58.

Vaughan, M. (1989). Rule-governed behaviour in behaviour analysis: A theoretical and experimental history. In: S. C. Hayes (Ed.), *Rule-governed behaviour: Cognition, Contingencies and Instructional Control* (pp. 97-115). New York: Plenum Press.

Weiner, H. (1969). Controlling human fixed interval performance. *Journal of the Experimental Analysis of Behaviour, 12,* 349-373.

Weiss, A.R. (1977). A behavioural approach to the treatment of adolescent obesity. *Behaviour Therapy, 8,* 720-726.

Wilson, G.T. & Fairburn, C.G. (1993). Cognitive treatments for eating disorders. *Journal of Consulting and Clinical Psychology, 61,* 261-269.

Yankura, J. & Dryden, W. (Eds.)(1997). *Using REBT with common psychological problems: A therapist's casebook.* New York: Springer Publishing.

Zarb, J. (1992). *Cognitive-behavioral assessment and therapy with adolescents.* New York: Brunner/Mazel.

Zettle R. (1990). Rule-governed behavior: A radical behavioral answer to the cognitive challenge. *Psychological Record. 40,* 41-49.

Zettle, R.D. & Hayes, S.C (1979). *The effect of social context on the impact of coping self-statements.* Unpublished manuscript, University of North Carolina, Greensboro.

Zettle, R.D. & Hayes, S.C. (1982). Rule-governed behaviour: A potential theoretical framework for cognitive-behavioural therapy. *Advances in Cognitive-Behavioural Research and Therapy, 1,* 75-118.