Bangor University

DOCTOR OF PHILOSOPHY

An experimental analysis of the momentary DRO schedule

Miller, Beth Y.

Award date:
1995

Awarding institution:
Bangor University

Link to publication

An experimental analysis of the momentary DRO schedule

Beth Y. Miller
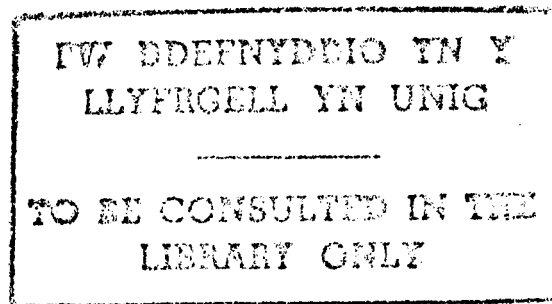
A thesis submitted in fulfilment of the regulations for the degree of Doctor of Philosophy in the University of Wales; 1996.

ACKNOWLEDGEMENTS

This thesis is dedicated to Lillie Crowther

SUMMARY

Whole-interval Differential Reinforcement of Other Behaviours (DRO), which reinforces set periods of behaviour absence, has been a widely used treatment in the reduction of inappropriate behaviours. Momentary DRO, a variant of this schedule which reinforces behaviour absence only at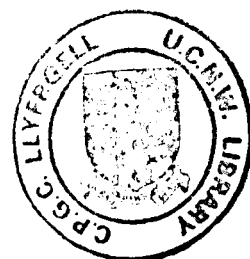 specific moments, may offer higher levels of reinforcement for clients and greater ease of administration. The primary aim of this thesis was to investigate the relative effectiveness of momentary DRO and whole-interval DRO in reducing inappropriate behaviour. The thesis also examined the claims that momentary DRO delivers greater reinforcement than whole-interval DRO and that momentary DRO is easier to programme.

Experiments 1 to 6 showed momentary DRO to be at least as effective as whole-interval DRO in reducing inappropriate behaviour. There was also evidence to suggest momentary DRO produced a more rapid behaviour reduction and that it was more effective in reducing non-targeted inappropriate behaviours. Experiments 1, 6 and 7 showed that momentary DRO consistently delivered higher levels of reinforcement than whole-interval DRO. Contrary to predictions, whole-interval DRO was reported to be as easy to administer as momentary DRO (Experiments 9 and 10). However, as subjects did not programme both schedules this result must be considered provisional. Experiments 9 and 10 showed that reinforcement was delivered more accurately in whole-interval DRO, although there was evidence to suggest that administration of this schedule was more effected by concurrent tasks than momentary DRO.

It can therefore be concluded that the effectiveness of momentary DRO has been under-estimated. It is clear, in view of its effectiveness and of the greater density of reinforcement it provides, that momentary DRO presents a viable, and in some cases superior, alternative to whole-interval DRO in the reduction of inappropriate behaviour.

CONTENTS

Chapter 1

The differential reinforcement of other behaviour schedule

"...determining the most effective means of applying differential reinforcement should be seen as a primary goal of applied behavior analysis in the treatment of behavior disorders" (Vollmer & Iwata, 1992, p. 394).

DIFFERENTIAL REINFORCEMENT

Differential reinforcement techniques are the most widely used treatments for reducing inappropriate behaviour in people with learning disabilities (Didden, Duker, & Korzilius, in press; Lennox, Miltenberger, Spengler, & Erfanian, 1988). They are all based on the principle that reinforcement is given for some behaviours (or rates of behaviour) and not others. There are four main types of differential reinforcement schedule, which are as follows:

- Differential reinforcement of other behaviour (DRO)
- Differential reinforcement of low rates of responding (DRL)
- Differential reinforcement of alternative behaviour (DRA)
- Differential reinforcement of incompatible behaviour (DRI)

DRO is the most widely used of the four techniques (Mazaleski, Iwata, Vollmer, Zarcone, & Smith, 1993), and its popularity is one reason why this thesis was conceived to investigate its variations and parameters. Before turning to DRO, however, the other three schedules will be briefly examined.

Differential reinforcement of low rates of responding (DRL)

DRL was first described by Skinner in 1938 and analysed further by Ferster and Skinner in 1957. In DRL, reinforcement is awarded for a particular inappropriate behaviour, contingent upon that behaviour occurring at a pre-determined low rate. DRL is useful when it is desirable to reduce, but not eliminate, a particular behaviour. DRL has been successfully used to reduce classroom disruption (Deitz & Repp, 1973; Deitz & Repp, 1974) and rapid eating (Favell, McGimsey, & Jones, 1980; Lennox, Miltenberger, & Donnelly, 1987).

Differential reinforcement of alternative behaviour (DRA)

This schedule entails reinforcing a specific and appropriate behaviour other than the one targeted for reduction (Deitz & Repp, 1983). A few studies have been reported which used DRA to reduce inappropriate behaviour (e.g. Young & Wincze, 1974; Saloviita, 1988), but in a review of twenty years of differential reinforcement procedures, O'Brien and Repp (1990) found so few examples of DRA (and DRL) that they discounted these from the rest of their analysis. In contrast, DRI and DRO were widely used.

Differential reinforcement of incompatible behaviour (DRI)

DRI is identical to DRA except that the behaviour which is chosen for reinforcement must be incompatible with the behaviour targeted for reduction. DRI has been shown to be successful in reducing stereotyped behaviour (Jones & Baker, 1988b; McClure, Moss, McPeters, & Kirkpatrick, 1986), pica (Smith, 1987), aggression (Friman, Barnard, Altman, & Wolf, 1986) and self-injurious behaviour (Tarpley & Schroeder, 1979). A number of reviewers (e.g. Cooper, 1987; Jones, 1991a; LaGrow & Repp, 1984) have noted that although DRI can be effective when used alone, it may be more effective when combined with other procedures. Studies which have successfully combined DRI with other treatments include Young and Wincze (1974), Azrin, Kaplan, and Foxx (1973), Azrin and Wesolowski (1980), and McGreevy and Arthur (1987). Generally the other procedures which are combined with DRI tend to be more punitive, such as overcorrection

and physical restraint. There have also been several studies which have not found DRI to be entirely successful in reducing stereotyped behaviour (e.g. Denny, 1980) or self-injurious behaviour (e.g. Borreson, 1980). Part of the problem, which is also true for DRO, is that some studies have not followed what are now regarded as standard procedures, such as choosing a potent reinforcer. This is considered in more detail below.

Differential reinforcement of other behaviour (DRO)

In DRO, reinforcement is provided if the target behaviour has not occurred for a specified period of time. All behaviours other than the target behaviour are effectively reinforced, hence the 'other' part of the name. This schedule was first described by Reynolds (1961), who was examining other aspects of responding and investigating why some behaviours increase as others decrease. He discovered that if he reinforced pigeons for not pecking after a period in which they had been reinforced for pecking, rates of this behaviour fell to a low level.

There are two types of DRO: whole-interval DRO and momentary DRO. In whole-interval DRO the subject has to refrain from the behaviour targeted for reduction for an entire interval of time in order to earn a reinforcer. In momentary DRO the subject has only to refrain from the target behaviour at the moment between two intervals. All the studies which are cited below used whole-interval DRO; momentary DRO has been very little used and is discussed in a separate section. Whole-interval DRO can be divided into a number of variations.

(i) Resetting/ Non-resetting DRO

In a resetting schedule, if the subject displays the target behaviour during an interval, then the interval is reset. Timing does not recommence until the subject has ceased the target behaviour. Then a new interval begins. Most uses of DRO employ a resetting feature (Jones, Walsh, & Sturmey, 1995; Vollmer & Iwata, 1992). In a non-resetting schedule, if

the target behaviour is emitted during the interval, the timing does not cease, but a reinforcer is not provided at the end of that interval, and then the next one begins.

(ii) Fixed-interval or changing-interval DRO

In fixed-interval DRO the interval during which the target behaviour must not be displayed remains the same throughout. This variation is quite widely used, but there are also two changing-interval procedures. The first is variable-interval DRO, which uses intervals of varying length, based round a mean length of time. For example, the standard DRO interval might be 30 seconds. The intervals used could then be 15 seconds, 45 seconds, 25 seconds, and so on. The second type of changing-interval DRO schedule is escalating DRO. Here, the DRO interval is static until the subject's behaviour is consistently reduced to a pre-determined criterion. Once this is reached, the interval is increased slightly, and is implemented until criterion is reached at that level. Then it is slightly increased once more (e.g. Cowdery, Iwata, & Pace, 1990).

Applications of DRO

The potentially useful human applications of Reynold's original procedure were quickly noted. Early studies included Patterson, Jones, Whittier, and Wright (1965) who used DRO to reduce hyperactivity in a child with learning disabilities, and Allen and Harris (1966) who used the technique to reduce a child's scratching. Since these early applications DRO has been increasingly employed as a reductive technique. In a wide-ranging review of treatments for inappropriate behaviour, Didden et al. (in press) found there to be twice as many studies involving DRO than the next most popular treatments (DRI, restraint and guided movement training). Lennox et al. (1988) found a similar trend in their review: amongst what they termed 'Level I' procedures (least intrusive interventions), differential reinforcement techniques were used with over twice as many subjects as other procedures. They did not, however, distinguish between DRO, DRI and DRL, but findings from other reviews would suggest that DRO made up the majority of differential reinforcement studies.

DRO has been used successfully to reduce stereotyped behaviour (Barton, Brulle, & Repp, 1986; Haring, Breen, Pitts-Conway, & Gaylord-Ross, 1986; Kennedy & Haring, 1993; Repp, Deitz, & Speir, 1974), self-injurious behaviour (Cowdery et al., 1990; Repp & Deitz, 1974; Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993), inappropriate masturbation (Foxx, McMorrow, Fenlon, & Bittle, 1986), aggression (Andrews, 1988; Frankel, Moss, Schofield, & Simmons, 1976; Redmon, 1987; Whitaker, 1992), disruptive behaviour (Repp, Barton, & Brulle, 1983; Repp, Deitz, & Deitz, 1976), vocal tics (Wagaman, Miltenberger, & Williams, 1995), thumb-sucking in non-learning disabled children (Christensen & Sanders, 1987; Lowitz & Suib, 1978), and smoking in a non-learning disabled adult (Barton & Barton, 1978).

DRO has been successfully used alone (e.g. Cowdery et al., 1990; Foxx et al, 1986; Haring et al., 1986; Redmon, 1987; Repp et al., 1983; Repp et al., 1976), and in combination with other procedures such as verbal punishment (e.g. Matson & Keyes, 1990; Poling, Miller, Nelson, & Ryan, 1978; Repp & Deitz, 1974), time-out (Peterson & Peterson, 1968), overcorrection (Azrin, Gottlieb, Wesolowski, & Rahn, 1975; Freeman, Graham, & Ritvo, 1975), extinction (Anderson, Dancis, & Alpert, 1978; Rubin, Griswald, Smith, & DeLeonarda, 1972) and contingent exercise (Luce & Hall, 1981). It has mostly been used with people with learning disabilities, but it has occasionally been used with non-disabled individuals (e.g. Barton & Barton, 1978; Christensen & Sanders, 1987; Leitenberg, Burchard, Burchard, Fuller, & Lysaght, 1977).

In some cases DRO has been shown to be more effective than procedures with which it is compared. For example, it has been found to be more effective than extinction (Favell, McGimsey & Jones, 1978; Goetz, Holmberg & LeBlanc, 1975; Luiselli, Helfen, Colozzi, Donellon, & Pemberton, 1978), time-out (Frankel et al., 1976; Repp & Deitz, 1974), verbal punishment (Repp & Deitz, 1974; Repp, Deitz & Speir, 1974) and medication (Luiselli, 1986).

Not all studies have reported success with DRO, however. Both Jones and Baker (1988a) and Tarpley and Schroeder (1979) found DRI to be superior to DRO in reducing stereotyped behaviour, and Friman et al. (1986) found DRI to be superior to DRO in reducing aggression. Barrett, Matson, Shapiro, and Ollendick (1981) and McGonigle and Rojahn (1989) found visual screening to be more effective than DRO in reducing stereotyped behaviour. Foxx and Azrin (1973) reported that of five treatments to reduce stereotyped behaviour, DRO was the least effective. Corte, Wolf, and Locke (1971) found DRO failed to reduce self-injurious behaviour. A few studies have reported that DRO actually increased the target behaviour (Derwas & Jones, 1993; Foxx & Azrin, 1973; Friman et al., 1986; Jones & Baker, 1988a).

Despite the large number of studies using DRO, a conclusive statement about its effectiveness cannot yet be made. Although a majority of published studies show DRO to be effective, it is not clear how many studies have not been published because they failed to show any improvement.

Theoretical underpinnings

Learning theory (e.g. Skinner, 1953) asserts that behaviours are acquired, developed and maintained by environmental events which can be reinforcing or punishing. For example, a person may engage in self-injurious behaviour to get attention from care-staff because s/he has found in the past that this behaviour reliably produces such attention (thus the behaviour provides positive reinforcement). Alternatively, an individual may have found that engaging in self-injury reliably removes them from an unpleasant or unwanted situation (negative reinforcement). Before Skinnerian conditioning principles were applied in the early 1960s, the prevailing belief was that people with severe learning disabilities were not capable of changing or learning (see Remington & Evans, 1988). The huge amount of research that has been carried out during the last thirty years has demonstrated quite the opposite. Indeed, Skinner's behavioural principles have probably nowhere had such success and durability than in the behaviour modification of people with learning

disabilities (Clarke & Clarke, 1987; Hergenhahn & Olson, 1993; Martin & Pear, 1996). Learning theory has led to a widely applied behavioural approach to inappropriate behaviour, where the emphasis is largely placed on the overt behaviour of the client, and how that behaviour alters in response to changes in the environment.

In addition to classic explanations of reinforcement and punishment, learning theory can provide a number of related reasons why inappropriate behaviour might arise and/or be maintained. These include the loss of usual reinforcers and a change of environment. Also relevant are the setting conditions outlined by behavioural theorists (Bijou, 1976; Kantor, 1959; Wahler & Fox, 1981) such as spatial density (Thelen, 1980; Warren & Burns, 1970), noise (Forehand & Baumeister, 1970; Jones & Carter, 1991; Levitt & Kaufman, 1965) and food (Miller, 1991; O'Banion, Armstrong, Cummings, & Stange, 1978).

Skinner's original ideas have been expanded upon by many authors. For example, the work of Iwata, Dorsey, Slifer, Bauman, and Richman (1982), Carr and Durand (1985a) and Durand (1990), has concentrated on the potentially communicative functions of inappropriate behaviour. Interest in the suggestion of inappropriate behaviour as a form of communication grew in the early 1980s (e.g. Donnellan, Mirenda, Mesraos, & Fassbender, 1984; Durand, 1982; Schuler & Goetz, 1981). Building on this work, Iwata et al. (1982) and Durand and Crimmins (1988) argued that it is possible to analyse the functions of such behaviour. The process by which this is carried out is termed functional analysis. Essentially, functional analysis describes any method which aims to determine the reasons why an individual continues to display a particular behaviour. Functional analysis was defined by Owens and Ashcroft (1982) as follows: "A functional analysis approaches a problem or a phenomenon seeking to answer questions regarding the function of the phenomenon to the system as a whole and the form of the relationship between this phenomenon and the function(s) it serves" (p. 188).

An individual may try and communicate their needs for attention, or for a particular object, or their wish to avoid a situation, via their inappropriate behaviour. Carr and Durand (1985b) advocated a procedure called functional communication training, whereby inappropriate behaviours are replaced with more appropriate ones. There is not yet sufficient evidence to state categorically that inappropriate behaviour is a form of communication, and Durand and Berotti (1991) stressed that communication should be regarded only as a metaphor. Not all inappropriate behaviours can be interpreted in a communication framework; some behaviours appear to be maintained by the stimulation they provide. That is, the behaviour is carried out simply because it feels enjoyable or provides needed stimulation. Behaviours like these, which are operantly maintained - but by factors other than those in the social environment - are described as being under the control of automatic reinforcement (Vaughan & Michael, 1982; Vollmer, 1994). A similar type of function has been described by Lovaas, Newsom & Hickman (1987) as perceptual reinforcement.

Functional analysis as a first step in devising treatment strategies is now regarded as standard practice (Jones et al., 1995). Nonetheless, many researchers do not report having used any method of functional analysis. In their review of treatments, Didden et al. (in press) reported that only 20.95% of studies carried out any pre-treatment functional analysis, which is even less than the 36% reported earlier by Lennox et al. (1988). This is surprising as there is a large body of evidence which indicates that a functional analysis can greatly enhance treatment effectiveness (Iwata et al., 1982; Repp, Felce & Barton, 1988; Vollmer et al., 1993).

However, there are some problems with functional analysis procedures. For example, it has been shown (Crawford, Brockel, & Schauss, 1992) that different methods of functional analysis can produce different outcomes. A number of authors have recommended that a combination of different analyses are carried out and that the variables which are proposed by these are tested experimentally (Crawford et al., 1992; Lennox &

-Miltenberger, 1989; Mace, Lalli, & Lalli, 1991). Another problem is that even if a function can be reliably determined under one circumstance, it is possible that the same behaviour may have a different function under other circumstances, depending on the past history of reinforcement of the subject. There is little scope in standard functional analyses for the possibility that behaviour may be multiply determined (e.g. Repp, Karsh, Deitz, & Singh, 1992), even though there is rarely a single 'cause' (Cullen, 1983).

These problems mean that the function of a person's inappropriate behaviour can be extremely difficult to determine (Hoyert & Zeiler, 1995). Nonetheless, the recent renewal of interest in functional analysis (Iwata & Fisher, 1994; Sturmey, 1996) suggests that there is a will to overcome these difficulties. Vollmer and Iwata (1992) predicted that the technology will improve, and they suggested that this will have wide implications for DRO procedures: "...as analyses of behavioral function become more common and more precise, the effectiveness of reinforcement-based procedures will increase" (p. 394). Some of the most widely used techniques of functional analysis are discussed in the next section.

FACTORS CONCERNING THE EFFECTIVENESS OF DRO

Most researchers accept that there are a number of measures which should be taken in order to maximise the effectiveness of a DRO schedule. Many of the published studies which show DRO to be unsuccessful have failed to take one or more of these factors into account. The factors are:

- implementing a functional analysis technique to try and ascertain the function of the behaviour
- choosing a potent reinforcer
- using an appropriate DRO interval

Implementing functional analysis

There are a number of methods which may be used, alone or in combination, which can provide some measure of behavioural function. The most widely cited method (Didden et al., in press) is the analogue assessment outlined by Iwata et al. (1982), which manipulates antecedent events and assesses their effect on behaviour. In their original demonstration of this technique, Iwata et al. (1982) presented a series of four environmental conditions to their subjects, termed 'alone', 'social disapproval', 'demand' and 'play', and recorded the rate of inappropriate behaviour in each. Analogue assessment is one of the most empirical methods of functional analysis. However, its implementation can be quite lengthy, because it requires the practitioner to set up several different conditions in which the client is placed and observed. One of the reasons why functional analysis is little used (as mentioned above) may be that it is perceived to be difficult and time-consuming (Ager, 1991; Reed & Head, 1993), and it is probable that analogue assessment is responsible for that impression. However, there are a number of other methods which are able to provide information about likely function and do not require a large expenditure of time.

For example, ABC charts have been quite widely used. These require care-staff to record occurrences of the behaviour, and note its antecedents and consequences. Observational methods such as scatterplots and naturalistic observation are also reported techniques of functional analysis, as are interviews with parents and care-staff. In addition, there are a number of checklists which can be completed by care-staff or parents. These include the Motivation Analysis Rating Scale (Wieseler, Hanson, Chamberlain, & Thompson, 1985), and the Functional Analysis Checklist (Van Houten & Rolider, 1991). However, the most widely used checklist is the Motivation Assessment Scale (Durand & Crimmins, 1988), which is a sixteen-item questionnaire. It aims to identify which one of four variables (sensory, escape, attention or tangible) is most likely to be maintaining the inappropriate behaviour. Although some studies have found the MAS to be unreliable (Newton & Sturmey, 1991; Zarcone, Rodgers & Iwata, 1991), others (Durand & Crimmins, 1988; Kearney, 1994) have found significant levels of reliability. Kearney recommended that

several raters should be asked to complete the MAS, as this would allow a more accurate assessment than just one individual rater.

It is not known how many DRO failures can be attributed to a lack of functional analysis, but it is likely to be implicated in a number of studies (Jones, 1991a). The main reason why functional analysis is important in DRO is because it allows an appropriate reinforcer to be chosen. This issue is considered in more detail in the following section.

The reinforcer

"Perhaps the greatest threat to the effectiveness of the...DRO procedure is the use of a stimulus that is only assumed to function as a reinforcer." (Cooper, 1987, p.394)

In examining the reinforcement component of DRO, Mazaleski et al. (1993) described two types of reinforcement: (1) the reinforcer which maintains the behaviour and (2) arbitrary reinforcers (e.g. any stimulus other than that which maintains the behaviour).

Before techniques of functional analysis were well-established, arbitrary reinforcers were the only ones used (Vollmer et al., 1993). Indeed, for quite some time researchers appeared to pay no attention at all to the reinforcer that they chose. A stimulus can only be termed a reinforcer if it can be shown to increase subsequent behaviour; but many studies failed to test their stimuli in any way. Often it was simply assumed that something would be a reinforcer, such as sweets and praise (see Jones, 1991a; O'Brien & Repp, 1990).

Reinforcers which have been arbitrarily chosen will only help to reduce an inappropriate behaviour if they are stronger than the reinforcers maintaining that behaviour (Vollmer & Iwata, 1992). Behaviours such as stereotypy tend to occur at a high rate; thus it can be assumed that whatever is reinforcing the behaviour, this reinforcer is occurring at a high rate also (O'Brien & Repp, 1990). The reinforcer chosen to reduce the behaviour must therefore be very strong and/or must be delivered on a rich schedule (i.e. very frequently).

Vollmer and Iwata (1992) recommended that at the very least, stimuli are chosen which have been shown to operate as reinforcers in another context. "By determining which stimuli function as effective reinforcers in one context, a degree of arbitrariness is eliminated when those stimuli are used in a differential reinforcement procedure" (p. 401). If non-functional reinforcers are to have any chance of success, then it is important that they are systematically tested in some way; this will at least give arbitrary reinforcers a chance to work (Mazaleski et al., 1993).

There are a number of methods which various authors have recommended for the selection of reinforcers (Berg & Wacker, 1991; Remington & Evans, 1988). For example, a reinforcer can be tested by being applied to a simple task, such as putting marbles in a box (e.g. Cooper, 1987). After recording baseline performance the reinforcer (or a variety of reinforcers) can be introduced; any concomitant increases in performance can be ascribed to the reinforcer. Another technique is to present subjects with a variety of items and record which are the most frequently approached (e.g. Pace, Ivancic, Edwards, Iwata, & Page, 1985).

Such methods can be an effective means of choosing a reinforcer (e.g. Steege, Wacker, Berg, Cigrand & Cooper, 1989). However, they do not always work. For example, a subject may consistently approach food in a reinforcer trial, then, during the treatment itself the subject may be sated and thus food will no longer be reinforcing. Or food may be reinforcing at the start of the session, but high levels of reinforcement may mean that satiation occurs. Alternatively, a stimulus may be consistently approached in a reinforcer trial, but it may not be strong enough to compete with the reinforcer maintaining the target behaviour (Mazaleski et al., 1993). It is important then that not only is the individual reinforcer considered but the setting condition (Kantor, 1959; Wahler & Fox, 1981) is also taken into account. One way round these problems is to identify and use a variety of reinforcers so that satiation and boredom are less likely to arise (e.g. Dyer, 1987; Egel, 1981).

Following from the work on functional analysis of Iwata and others, researchers have realised that for maximum effectiveness the reinforcer should be of a similar type to that maintaining the inappropriate behaviour. Vollmer and Iwata (1992) suggest that, "...when differential reinforcement procedures are prescribed by a functional analysis, the reinforcers used are both functional and nonarbitrary" (p.402).

However, even if the function of a behaviour can be determined, appropriate reinforcers must still be found. For example, a person's self-injury may be maintained by attention, but this does not mean that all types of attention would be reinforcing. They may, for instance, only respond to praise, or they may only respond to attention from one person. Therefore, reinforcer assessments should still be carried out.

It does not appear that many researchers have used a functional analysis approach to reinforcer selection, or have even attempted to assess the reinforcing qualities of arbitrary reinforcers. O'Brien and Repp (1990) reported in their review of DRO that in 78% to 98% of studies there was no indication as to whether a reinforcer had been tested for its reinforcing properties before the study took place. They found that edibles were by far the most widely used reinforcer, with social reinforcers second, yet these were the two least effective reinforcers. Social reinforcers were particularly ineffective, but this may be in part because they were used quite widely with more extreme inappropriate behaviours such as aggression and self-injury. The most effective reinforcers were restraint, vibration and olfactory stimulation. However, caution is necessary as the sample sizes of these studies were small.

Choosing potent reinforcers for DRO schedules is not a new idea. As early as 1974, Repp and Deitz (1974) stated that for DRO to be effective it would need "...the identification of reinforcers stronger than those maintaining the maladaptive behavior" (p. 324). Reinforcement is one factor which may well have contributed to the success or otherwise of DRO programmes. Poling and Ryan (1982) suggested that "...when the DRO fails to

control behavior, the first possibility to consider is that the stimulus is not a sufficiently powerful reinforcer" (p.16). Mazaelski et al. (1993) reported that two of the studies which showed a lack of success with DRO (Corte et al., 1971; Foxx & Azrin, 1973) failed to use reinforcers of the type maintaining their subjects' inappropriate behaviour. Repp et al. (1983) are widely quoted when they say that unsuccessful DRO studies "...may have demonstrated the extent to which some researchers are poor or unlucky at the task of selecting reinforcers, rather than the extent to which some schedules are ineffective" (p. 444).

The DRO interval

An appropriate interval length is another important element which can influence the success of DRO. In DRO the subject is required to refrain from the target behaviour for a certain period of time in order to receive reinforcement: this period is the DRO interval. Repp, Deitz, and Speir (1974) proposed a method for calculating the interval which appears to have been generally accepted since. This uses the inverse of the mean response rate during baseline. In other words, if the subject emitted the target response six times a minute on average during baseline, then the DRO interval would be one minute divided by six - that is, ten seconds. Repp et al. (1974) suggested that by using this method, the rate of reinforcement available for not responding would be equal to that available for responding. An alternative calculation records the inter-response times during baseline (the length of time between occurrences of the behaviour) and then calculates the mean.

Some researchers have advocated setting the DRO interval length to be slightly shorter than the mean inter-response time (e.g. Jones, 1991a) to increase the possibility of the behaviour contacting the contingencies. Repp and Slack (1977) showed that DRO was more effective when the interval was small than when it was large (at least initially). They suggested that this single variable could be responsible for the success or failure of a DRO intervention. Similarly, Repp, Felce and Barton (1991) showed that using a DRO interval which was equal to the mean rate during baseline was more effective than when it was

twice the mean rate. Essentially, if an interval is too small, then satiation might occur, and if it is too large there is little chance of the subject refraining from the target behaviour long enough to earn reinforcement (O'Brien & Repp, 1990).

Many researchers propose that the DRO interval should be increased as the subject performs to criterion (Poling & Ryan, 1982; Repp & Slack, 1977; Vollmer & Iwata, 1992). This is sometimes termed escalating DRO. In their review of reinforcement-based procedures O'Brien and Repp (1990) reported that none of the articles provided any information as to whether they had in fact carried out such a procedure. Nonetheless, some studies have employed an escalating DRO schedule (Cowdery et al., 1990; Poling et al., 1978; Repp & Deitz, 1974; Repp et al., 1974; Repp & Slack, 1977; Topping, Graves, & Moss, 1975).

Not all studies have based their DRO interval on the behaviour as it occurred in baseline. Jones (1991a) cited the example of Borreson (1980), who used an interval of five minutes, despite the fact that at baseline the self-injurious behaviour occurred approximately six times a minute. Using the procedure outlined by Repp et al. (1974) the DRO interval should in fact have been ten seconds or less. The DRO schedule did not reduce the subject's behaviour, presumably because it did not contact the contingencies at any point. Repp et al. (1983), despite being early advocates of relevant intervals, did not base their DRO interval length on behaviour during baseline. They used the same DRO interval of five minutes for each of their three subjects. In this case, however, they found DRO to be an effective procedure.

It is possible that some of the reasons for the mixed findings on DRO have been due in part to the interval length.

In addition to factors which may influence the effectiveness of DRO, there are also two practical considerations which should be taken into account before a DRO schedule is used. These are the generalisability of DRO, and its side-effects. These are briefly considered below.

Generalisation

It is not clear how well DRO generalises over time or to novel settings (Homer & Peterson, 1980). The animal literature has shown that DRO produces longer lasting effects than DRA, extinction and punishment (e.g. Mulick, Leitenberg, & Rawson, 1976; Uhl & Garcia, 1969; Uhl & Sherman, 1971). Few studies, however, have looked at the relative generalisability of treatments in humans. It appears that the reduction of behaviour produced by DRO typically ends when the DRO contingencies are removed (Hoyert & Zeiler, 1995; Poling & Ryan, 1982). This was shown by Jones, Baker, and Murphy (1988), who found reductions in behaviour were not maintained beyond the DRO intervention. A few exceptions have been noted. Homer and Peterson (1980) cited an unpublished study which showed DRO to produce a more durable reduction of self-injurious behaviour than electric shock. Barton and Barton (1978) showed that a DRO-produced elimination of smoking was still in effect thirty months later; significantly it appeared that the reduction had come under the control of natural contingencies such as peer approval and health benefits. There have been informal reports which suggest that DRO can generalise to other settings (e.g. Peterson & Peterson, 1968), and a few studies which have made generalisation part of the treatment goal (e.g. Christensen & Sanders, 1987).

There is some evidence that unless generalisation is programmed into a DRO schedule, DRO is not particularly durable. This evidence comes from the many reversal designs which have often shown DRO to revert to baseline levels immediately on return to baseline

(e.g. Poling et al., 1978; Repp, Deitz & Deitz, 1976; Repp, Deitz & Speir, 1974). Hoyert and Zeiler (1995) have recently outlined procedures which can be added to DRO to enhance its durability. However, their work was with pigeons; the generalisability of DRO with humans remains largely unexplored.

Side effects

For some time it was thought that DRO did not produce any undesirable side effects. For example, Homer and Peterson (1980) concluded that any side-effects of DRO were likely to be positive (such as strengthening desirable behaviours). However, they conceded that more research was necessary.

Cowdery et al. (1990) researched the possible side-effects of DRO. They carried out an escalating DRO schedule with a nine year-old boy who scratched himself. Token reinforcement was used in the form of pennies which could be exchanged later for treats. They began with a DRO interval of two minutes, and this was gradually increased throughout the study until it reached thirty minutes. The DRO schedule itself was highly effective in reducing the boy's self-injurious behaviour. However, the authors noted that after the first DRO session in which the boy scratched himself and was told he would not receive a penny, he began to cry. Thereafter, crying or shouting tended to occur whenever he displayed self-injury during a DRO interval, presumably because he knew then that he would not receive a penny. This study showed that DRO is not necessarily an entirely positive procedure: it can be associated with negative emotions in the same way as punishment can. As Cowdery et al. put it, "...occasions where reinforcement is not delivered have the potential for generating behavior that typically is associated with deprivation states, reinforcement withdrawal, or aversive stimulation" (1990, p. 505). It may be worth noting, however, that Cowdery et al. did inform the boy of the contingencies in operation. Although this is generally considered to be good practice, it is possible to speculate that similar side-effects would not have been noted had the boy been unaware of

the contingencies. Whether or not a reduction in self-injury would still have ensued is an empirical question which could be addressed by future research.

Vollmer et al. (1993) suggested that the negative emotional side-effects shown by Cowdery et al. (1990) may have been due to the extinction component of the DRO schedule. Vollmer et al. (1993) themselves found some evidence of negative side-effects of DRO; when compared to non-contingent reinforcement, DRO resulted in greater occurrences of aggression and disruption, although these disappeared before the end of treatment.

Rolider and Van Houten (1990) proposed that DRO, leading as it does to the reduction of reinforcement following a response, is actually a negative punishment procedure. Presumably if this is the case, then DRO would be as prone to negative side-effects as other punishment techniques. This viewpoint is somewhat controversial, however (Repp & Singh, 1990).

The potential side-effect of DRO which has been most widely discussed (e.g. Repp & Deitz, 1974) is that of adventitious or superstitious conditioning (Skinner, 1948): that is, inadvertently reinforcing another behaviour such that it increases beyond baseline levels. Zeiler (1970) demonstrated in an experimental situation that such adventitious strengthening was possible in a DRO schedule. His subjects were children who had to refrain from key-pressing in order to receive a sweet. He found that for two of his six subjects, the behaviour they were engaging in when a reinforcer was administered became strengthened (in the case of one boy this behaviour was searching on the floor for a dropped sweet). The implication is that DRO can inadvertently reinforce other inappropriate behaviours if they occur frequently enough to coincide with reinforcement provided for non-responding of target behaviour. However, this phenomenon has been observed with schedules other than DRO and so may be a feature of reinforcement schedules in general, rather than DRO specifically (Zeiler, 1970).

There is little recorded evidence in the literature that DRO does result in adventitious conditioning, but this may be because very few studies report the effects of the schedules on other behaviours. Informal evidence, however (e.g. Bostow & Bailey, 1969; Repp, Deitz, & Speir, 1974) suggests that adventitious conditioning of undesirable behaviour does not often occur.

To avoid inadvertently reinforcing other inappropriate behaviours, several authors have recommended that reinforcement should not be given shortly after *any* display of undesirable behaviour, even if the target behaviour has not occurred during that interval (Jones et al., 1995; Poling & Ryan, 1982; Repp & Deitz, 1979). Repp and Deitz (1979) suggested that a list of all possible undesirable behaviours for a particular client should be given to staff so that they know which behaviours should not coincide with a reinforcer.

It is also possible that other inappropriate behaviours may increase if the target behaviour is reduced, because of response covariation (e.g. Adlem & MacDonald, 1992; Singh & Repp, 1988). This refers to changes in the rate of behaviours other than the one targeted for reduction, which are influenced by the manipulation of the target behaviour (Jones & Baker, 1989). Response covariation has been seen in treatments such as overcorrection (Ollendick & Matson, 1978), punishment (Rollings & Baumeister, 1981) and DRL (Singh, Dawson, & Manning, 1981). Measures of response covariation are rarely taken in DRO studies, though they are sometimes informally reported (Repp & Deitz, 1974; Repp, Deitz, & Speir, 1974). Derwas and Jones (1993) did take formal measures of other inappropriate behaviours and found that for some of their subjects the other-behaviours increased as the targeted behaviour decreased. Horner and Barton (1980) and Jones et al. (1995) recommended that such measures are taken as a matter of course, so that the effect of DRO can be seen on non-targeted behaviours, including any positive alternative behaviours which may be strengthened.

The available evidence shows that negative side-effects such as emotional behaviour and adventitious conditioning occur only occasionally in DRO. However, there are two other problems which need to be considered in detail. These are outlined in the following section.

<div align="center">PROBLEMS WITH DRO</div>

Whole-interval DRO has been widely used to reduce a variety of different behaviours. However, it presents two main problems, which suggest the need to evaluate an effective alternative. These problems are (i) low levels of reinforcement and (ii) difficulty of use. These are discussed below.

Low levels of reinforcement

People with learning disabilities, particularly those living in institutions, often have very limited access to reinforcers and experience little social interaction (Mazaleski et al., 1993). For some researchers (e.g. Vollmer et al., 1993) the potentially low level of reinforcement provided by a DRO programme is problematic, as it exacerbates the deprived state of the individual.

To examine why whole-interval DRO tends to provide low levels of reinforcement, it is necessary to consider briefly what actually happens during a DRO procedure. A DRO schedule is made up of two components: a reinforcement component and an extinction component (Mazaleski et al., 1993). Reinforcement is provided if a certain period of time has passed without presentation of the target behaviour. Extinction occurs when reinforcement is withheld contingent on presentation of the target behaviour. The relative importance of these two components in the effectiveness of DRO has been examined by Mazaleski et al. (1993). They carried out a component analysis of DRO with three individuals who displayed socially-reinforced self-injurious behaviour. They found that it made little difference whether they employed an arbitrary or non-arbitrary reinforcer; the

single relevant factor in the reduction of the self-injury was the withdrawal of social reinforcement contingent on the target behaviour - extinction.

If the maintaining variable of the inappropriate behaviour is not correctly determined, then a DRO schedule that does not contain an extinction component will ensue (Mazaleski et al., 1993). For example, the maintaining variable for a person's self-injury might be attention, but if an arbitrary reinforcer such as food was used in a DRO schedule, and the care-worker continued to provide attention contingent on the self-injury, then this would not actually be a DRO schedule. In such a case as this there would be little likelihood of the target behaviour being reduced, and Mazaelski et al. suggested that this may account for some DRO failures. They concluded that the relevance of the reinforcer is vital, not for the reinforcement component to work, which is what others have tended to stress, but for the extinction component to work. They accepted that DRO schedules can work if an arbitrary reinforcer is used - as long as it is a strong reinforcer - but a durable and strong reduction is unlikely to occur without the extinction component. Both Mazaelski et al. and Vollmer et al. (1993) considered that the real task of the reinforcement component is to mitigate deprivation rather than reduce targeted behaviour. To combine the effectiveness of extinction with a richer source of reinforcement, they both recommended using DRO schedules which "...maximise access to reinforcers while minimizing reinforcers for undesirable behavior" (Mazaleski et al., 1993, p.155). Both suggested that momentary DRO would fulfil this requirement more successfully than whole-interval DRO.

Difficulty of use

As discussed in the preceding sections, whole-interval DRO has been shown in the literature to be generally successful in the reduction of inappropriate behaviour. However, it has had less success in applied settings (Ager, 1991; Oliver, Murphy, & Corbett, 1987), and this may be because the successful implementation of whole-interval DRO requires high levels of concentration and attention from the programmer (Cooper, 1987; Poling & Ryan, 1982). Tierney, McGuire, and Walton (1979) could not continue using a DRO

programme because staff reported so much difficulty in its implementation. Jones and Baker (1988b) and Saloviita (1988) encountered similar problems with the Differential Reinforcement of Incompatible behaviour (DRI) schedule, which like whole-interval DRO, requires continuous subject monitoring.

Although Repp and Deitz (1979) drew up guidelines intended to help staff run effective DRO programmes, they perhaps overestimated the amount of time, energy and motivation available to care-staff working in busy and stressful conditions. "If a profoundly retarded client responds to little less than a song and dance, then staff should learn to sing and dance" Repp and Deitz (p. 222) suggested, but this implies high staff numbers and freedom from daily chores that is not necessarily a feature of many institutions.

Many studies of treatment efficacy take place in highly-controlled environments with favourable experimenter:subject ratios (Jones, 1991a). The intensity necessary for effective programming is difficult outside of the research setting, where the member of staff operating the DRO schedule is likely to have a number of concurrent responsibilities. In a study which compared DRO with non-contingent reinforcement, Vollmer et al. (1993) stated that, "... differential reinforcement can be cumbersome to administer over long periods of time because it often requires continuous monitoring of a client's behaviour" (p. 10). They agreed with Jones (1991a) that it is care-staff with other responsibilities who will encounter particular difficulty with whole-interval DRO.

Poling and Ryan (1982) recommended that "future investigations of DRO... [should] address issues of practicality" (p.15). A few recent studies have examined the little-used variant of DRO, momentary DRO, which appears to be less time-consuming in its implementation than whole-interval DRO, and so may be of more use in applied settings (Jones et al., 1995). These studies are considered below.

MOMENTARY DRO

In momentary DRO, the reinforcer is presented if the target behaviour is absent at the *moment* between two intervals, rather than being absent for the entire interval (as in whole-interval DRO). Momentary DRO was first described by Sulzer-Azaroff and Mayer in 1977, although they appeared to believe that they were referring to the standard form of DRO. Whole-interval DRO was referred to as a "common variation" (p. 256). Very few studies have been carried out which have employed the momentary DRO procedure. One of the first was by Harris and Wolchik in 1979. They compared three interventions for the reduction of stereotypy in four subjects: time-out, momentary DRO and overcorrection. They found that overall, overcorrection was the most effective procedure, and momentary DRO was the least. Although it produced a small decrease in stereotypy in one subject, it had no effect with two of the others and was associated with an increase in stereotypy in the fourth. Mixed results were also found by Sisson, Van Hasselt, Hersen, and Aurand (1988), who found momentary DRO alone to be effective in reducing the inappropriate behaviours of one of three subjects. However, it was effective for the remaining two subjects when it was combined with time-out and overcorrection.

A few studies have compared the relative effectiveness of momentary DRO and whole-interval DRO, with varying results. Repp, Barton and Brulle (1983), who were the first to name momentary DRO, compared the two schedules' effectiveness in reducing disruptive classroom behaviour in three children. They found whole-interval DRO reduced disruption far more effectively than momentary DRO. In a second study they compared the schedules in the reduction of one boy's disruption, using a multiple-schedule design. Again, they found whole-interval DRO to be more effective than momentary DRO. The results of these studies led Repp et al. to suggest that while momentary DRO might not be strong enough to effect an initial reduction in target behaviour, it might be useful to continue suppression once whole-interval DRO had reduced the behaviour. They considered that this would be

easier for care-staff than to continue with a whole-interval DRO intervention, as "...the momentary DRO schedule is much easier to program" (p. 444).

The suggestion that momentary DRO could be used to continue the suppression begun by whole-interval DRO was tested by Barton, Brulle, and Repp (1986). They worked with nine subjects who displayed a variety of inappropriate behaviours, including different stereotypies and non-compliance. These nine were divided into three groups. The first group was presented with an ABC design which consisted of baseline, whole-interval DRO and momentary DRO. The second group, who were paired with the first group, were presented with an AB design, which was baseline and whole-interval DRO. This was to compare the effect of continuing with whole-interval DRO with moving onto momentary DRO. The third group were presented with an ABA(B) design, so that control of whole-interval DRO over the behaviour could be determined. This study confirmed that momentary DRO could be effective in continuing suppression of behaviour after the initial suppression had been carried out by whole-interval DRO.

These few studies appeared to demonstrate the relative inferiority of momentary DRO, so much so that Repp, Felce and Barton (1991) were able to state categorically that "...mDRO when used alone, is ineffective at reducing inappropriate behavior...Thus one variable associated with the effectiveness of DRO has been identified" (p. 418).

However, in 1993 Derwas and Jones compared the effectiveness of momentary DRO and whole-interval DRO in reducing high-rate stereotyped behaviour. They found that momentary DRO was more effective than whole-interval DRO for two of their subjects, as effective for two, and less effective for one. This study was the first to show that in some circumstances momentary DRO might be as or more effective than whole-interval DRO.

There were several methodological differences between these studies which may account for their discrepant findings. For example, Repp et al. (1983) made use of intervals of five

minutes - considerably longer than the majority of applied studies using DRO; and these intervals were not based on the rate of responding during baseline (Repp et al., 1974; Repp & Slack, 1977). Although this obviously did not have a bearing on the effectiveness of whole-interval DRO, it is possible that it worked to the detriment of momentary DRO. The rationale stated by Repp et al. (1983) for using identical five-minute DRO intervals for all three subjects was that it made it much easier to run the schedules concurrently in a classroom setting. The interval lengths used by Derwas and Jones (1993) were based on the rate of individual behaviour during baseline, and were much shorter (ranging from 15 to 30 seconds).

There were several differences between the subjects who participated in each of these two main comparison studies. Firstly, Repp et al. described the subjects in their first study as 'mildly retarded', with good language and self-help skills. The subject in their second study was 'moderately retarded', and also had self-help skills and quite good language. Derwas and Jones's subjects, on the other hand, were described as having 'severe mental retardation'; all were non-verbal and had very few self-help skills. Secondly, the behaviours Repp et al. aimed to reduce were disruptive classroom behaviours such as 'interruptions' and 'out of seat', whilst Derwas and Jones worked with stereotyped behaviour such as 'head-patting' and 'scrabbling'. Although the rate of the disruptive behaviour examined by Repp et al. was not stated, it is likely that it was of a lower rate than the stereotypy examined by Derwas and Jones, who reported that the mean inter-response times of their subjects ranged from 21 to 39 seconds.

Neither of these studies based their choice of reinforcer on hypotheses of the causes of the behaviours, and did not carry out any form of functional analysis (see Repp, Felce, & Barton, 1988). Informal observation from Derwas and Jones suggested that the stereotypy in all cases was self-stimulatory; Repp et al. made no suggestion of the functions of the behaviours. However, Repp et al. did assess the reinforcers they chose for their reinforcing potential; they tested different reinforcers on a simple task to assess which ones

were associated with higher responding. Derwas and Jones based their choice of reinforcers on recommendations of care-staff.

Perhaps the most salient difference between the two studies was the use of instructions. As their subjects were non-verbal, Derwas and Jones made no attempt to inform them of the contingencies in operation at any time, and the switch from whole-interval DRO to momentary DRO was not indicated in any way. In contrast, Repp et al. told their subjects in the whole-interval DRO condition that they would earn a reward if they were non-disruptive "for an entire 5 minutes" (p. 439). In the momentary DRO condition subjects were told, "If you are not being disruptive when Ms Smith raises her hand, you will earn a treat. If you are being disruptive, you will not earn a treat" (p. 438). In other words, Repp et al.'s subjects were told quite clearly that in whole-interval DRO they had to be good for the whole time, while in momentary DRO they were told that it was only at a particular moment that they had to be good. Perhaps then, it is not surprising that momentary DRO was ineffective in behaviour reduction. It cannot have taken the (fairly able) subjects long to realise that in momentary DRO they could continue disrupting and still earn a reinforcer, as long as they were quiet for a few seconds every five minutes. The fact that the DRO interval did not vary must have made it even easier for subjects to work out when the next reinforcement was due. Repp et al. were inadvertently guilty of what Homer and Peterson (1980) call 'strawman tactics'; that is, that in a comparison study they did not ensure that both procedures were operating at full efficiency. What Repp et al. compared was "...a procedure operating at full strength and a procedure with far less than optimal parameters" (Homer & Peterson, 1980, p. 456). Momentary DRO can probably never be fully effective if subjects are told that they need only behave appropriately at one particular moment.

The work by Repp et al. (1983) is one of the two main studies which provide evidence for the ineffectiveness of momentary DRO. The other is the earlier study by Harris and Wolchik (1979). As already mentioned, they found momentary DRO to be the least effective of three treatments. However, there were some flaws in Harris and Wolchik's

experiment. In particular, they used food and praise as reinforcers without any evidence that these were actually reinforcing for their subjects. Several recent studies (e.g. Green, Reid, White, Halford, Brittain, & Gardner, 1988; Pace et al., 1985) have shown that for some people with severe learning disabilities, social reinforcers such as praise are "...among the least preferred consequences" (O'Brien & Repp, 1990, p. 154).

Almost all of the studies which employed momentary DRO commented on its relative ease of use. Repp et al. (1983) called it "...the much easier DRO schedule to program" (p. 436) and Barton et al. (1986) stated that the use of momentary DRO would "reduce the amount of teacher time involved in carrying out the program" (p. 282). Sisson et al. (1988) highlighted the relative ease of training the effective implementation of momentary DRO to staff who were "...previously naive to the theories and techniques of behavioural intervention" (p.524). Derwas and Jones (1993) commented that their positive evaluation of momentary DRO will be useful "...for its application in residential settings since it is economical with staff time" (p. 53).

That momentary DRO appears to be easier to use than whole-interval DRO is a strong justification to investigate it further. As outlined above, another reason is the higher rate of reinforcement it might provide. Recently Vollmer et al. (1993) compared whole-interval DRO to non-contingent reinforcement (NCR) in the reduction of self-injurious behaviour. They defined the non-contingent procedure as "...a response-independent or time-based delivery of stimuli with known reinforcing properties" (p. 10). In other words, regardless of what the subject is doing at a designated moment, reinforcement is provided. NCR has several potential advantages over whole-interval DRO (Vollmer et al., 1993). Firstly, NCR may reduce extinction-related behaviour that is sometimes seen in DRO, because the functional reinforcer is still freely available. Secondly, NCR ensures that overall rates of reinforcement remain high, as, unlike in DRO, occurrences of the target behaviour do not make reinforcement less likely. In cases where the inappropriate behaviour is maintained by attention (as for the subjects in Vollmer et al.'s study), it would be undesirable to

deprive people who have limited means of obtaining attention. Vollmer et al. regarded the low levels of reinforcement in DRO as one of its main shortcomings: "...it is possible that some subjects would rarely receive social interaction if it were contingently delivered throughout the day" (p.18). Vollmer et al. were also concerned about the underlying message that DRO gives to staff: "A resetting DRO, if followed correctly, can validate or even mandate low frequencies of staff interaction and create further deprivation" (p. 10). Thirdly, NCR should be easier to use than whole-interval DRO.

Vollmer et al. (1993) found that NCR was as effective as DRO. The authors commented that the main negative side-effect that might be expected from a NCR procedure, adventitious conditioning of the target behaviour, did not occur. However, in such a procedure there is a very real possibility that some instances of the target behaviour would be reinforced, and so Vollmer et al. recommended using a combination of whole-interval DRO and NCR, which is in fact momentary DRO: "...a momentary time sample of behavior could be conducted on a fixed-time schedule, and if the subject was not engaging in SIB at that moment, attention could be delivered (see Repp, Barton, & Brulle, 1983)" (p.19).

Momentary DRO, then, has some features in common with non-contingent reinforcement. Like non-contingent reinforcement, it allows reinforcement to be delivered at a relatively high level, as unlike in whole-interval DRO, not all instances of the target behaviour remove the possibility of reinforcement. Non-contingent reinforcement may work because it combines satiation (a rich schedule of reinforcement) with extinction (removing the relationship between behaviour and reinforcer) (Vollmer et al., 1993). It may be that momentary DRO works in the same way, with the added benefit that no instances of the target behaviour could be inadvertently reinforced.

The conclusions of Vollmer et al. are echoed by Mazaleski et al. (1993). Having found that the extinction component is the active ingredient in DRO (at least where the functional

reinforcer is used), they suggested that the main purpose of the reinforcer is to avoid deprivation. Like Vollmer et al., they mentioned the "interesting procedure" used by Repp et al. (1983): "...momentary DRO involved time-based delivery of reinforcement (NCR), which was suspended if the target behavior was occurring at the scheduled time of delivery (extinction)" (Mazaleski et al., p.155). They concluded that even though Repp et al. (1983) found momentary DRO to be less effective, it "...seems even more attractive as an alternative to traditional DRO and deserves further investigation" (p. 155).

AIMS OF THE PRESENT THESIS

The literature in this chapter highlights two reasons why momentary DRO is worth closer investigation. Firstly, because it is generally considered to be easier to use than traditional whole-interval DRO, as it does not require continual monitoring of a client and constant resetting of a timing device. Secondly, because it combines the best elements of non-contingent reinforcement, such as high rates of reinforcement, with the best elements of DRO, such as never reinforcing the target behaviour.

If it transpires that momentary DRO can be as effective a reduction technique as whole-interval DRO, then the reasons described above would clearly make it the treatment of choice when behaviour reduction is necessary.

This thesis has three interrelated aims:

(1)   To examine the effectiveness of momentary DRO under a variety of different conditions.

(2)   To establish whether there is a higher density of reinforcement in momentary DRO than whole-interval DRO.

(3)   To investigate the relative ease of use of momentary DRO.

These aims are outlined in more detail below.

(1)   The effectiveness of momentary DRO

As outlined previously, two important variables are the type and rate of behaviour. Repp et al. (1983) worked with comparatively low-rate disruptive behaviour and found momentary DRO to be less effective than whole-interval DRO, while Derwas and Jones (1993) worked with high-rate stereotypy and found momentary DRO to be more effective than whole-interval DRO. The findings of these studies might suggest that momentary DRO would be less effective for disruptive or low-rate behaviours than whole-interval DRO. Early work

on DRO suggested that whole-interval DRO would be unlikely to succeed with low-rate behaviour (e.g. Homer & Peterson, 1980), but Whitaker (1992) showed that DRO could successfully reduce aggression which occurred only once or twice a month. It would be useful to know if momentary DRO could exert control over lower rate behaviours, and this variable is examined in two experiments which aim to reduce disruptive behaviour. The effectiveness of momentary DRO on high-rate stereotyped behaviour is also examined, in order to determine whether the findings of Derwas and Jones (1993) can be replicated.

This thesis also examines a hypothesis proposed by Repp et al. (1983). Their subjects had mild or moderate learning disabilities, and momentary DRO was not as effective as whole-interval DRO. The authors predicted that it would be even less effective with more capable subjects who could readily discriminate between the schedules. This variable is explored through some of the subjects who took part in these experiments, who either had mild learning disabilities, or no disabilities at all. It would be very useful to know whether momentary DRO is of use with non-disabled subjects such as school-children. Another related variable which may have a bearing on the effectiveness of momentary DRO is language. One of the main discrepancies between the studies of Repp et al. (1983) and Derwas and Jones (1993) was the use of instructions in the former study. This thesis examines the effects of working with verbal subjects on the results of the two schedules. If there are differences in schedule patterns this will have implications for choosing schedules to match clients' verbal ability.

Whole-interval DRO is flexible enough to have been used widely in conjunction with other treatments; often its effectiveness has been enhanced when it has been combined in this way (Cooper, 1987). It is not yet clear whether momentary DRO can perform a similar function. This, then, is the third variable concerning the effectiveness of momentary DRO this thesis examines: the contribution it may make to a treatment package. An experiment was carried out to assess whether momentary DRO could be used as part of a package to selectively reduce stereotyped behaviour. This may be a very important function for it to

fulfil, as current trends suggest that a total reduction of stereotypy is unnecessary and may be detrimental to the individual as it takes away a source of stimulation.

(2)   The density of reinforcement in momentary DRO

A number of authors have suggested that momentary DRO provides higher rates of reinforcement than whole-interval DRO (Mazaleski et al., 1993; Vollmer et al., 1993). This thesis aims to determine if there is experimental evidence for this. If this suggestion is supported, this thesis aims to explore whether momentary DRO is more effective than whole-interval DRO simply because it provides a richer schedule of reinforcement.

(3)   The ease of use of momentary DRO

As outlined elsewhere, momentary DRO is considered to be easier to programme (e.g. Repp et al., 1983; Barton et al., 1986), which may have a huge bearing on its efficacy in the 'real world'. The reason it is regarded as easier is because momentary DRO requires lower levels of concentration than whole-interval DRO. Because of this, it should be possible to carry out momentary DRO schedules while attending to other tasks. The assumption is that outside distractions will not interfere with momentary DRO because, unlike whole-interval DRO it does not require continual concentration on the part of the programmer.

This thesis aims to discover whether momentary DRO is in fact easier to programme than whole-interval DRO. If it transpires that it is easier, it might be expected that momentary DRO would also be programmed more accurately. Therefore, an exploration of the two schedules' accuracy of programming is carried out.

Chapter 2

The influence of the type and rate of behaviour, and the ability of the subject, on the relative effectiveness of momentary DRO and whole-interval DRO

The four experiments in this chapter explore the relative effectiveness of whole-interval DRO and momentary DRO with high-rate stereotyped behaviour in people with severe and profound learning disabilities, and with lower-rate disruptive behaviour in non-learning disabled and mildly disabled children.

As mentioned in the last chapter, two previous studies have compared the relative effectiveness of momentary DRO and whole-interval DRO. Repp et al. (1983) found whole-interval DRO to be the more effective schedule, while Derwas and Jones (1993) found momentary DRO was more effective. These studies differed from each other on three factors, which may have been responsible for their discrepant findings. The first factor on which they differed was the type of targeted behaviour. Repp et al. worked with disruptive behaviour, while Derwas and Jones worked with stereotyped behaviour. The second factor was the rate of behaviour. Stereotypy tends to occur at a high rate, and disruptive behaviour tends to occur at a comparatively lower rate. The discrepant findings of the two studies might suggest that momentary DRO is more effective for high-rate stereotyped behaviour and whole-interval DRO is more effective for lower-rate disruptive behaviour.

The third factor on which the two studies differed was the ability of the subjects who took part. Repp et al. worked with children who had mild-to-moderate learning disabilities and good language skills; Derwas and Jones worked with children and adults who had severe or profound learning disabilities and little or no language skills. Their findings might

suggest that momentary DRO is more effective with people who have profound or severe learning disabilities, and whole-interval DRO is more effective with people who have less severe disabilities.

The first two experiments in this chapter are developments of the Derwas and Jones (1993) study. The last two experiments are developments of the Repp et al. (1983) study.

The experiments in this chapter explore some of the aims outlined in 'Aims of the present thesis', as follows:

1.   To assess the effectiveness of momentary DRO in reducing high-rate stereotypy and lower-rate disruptive behaviour. (All experiments.)

2.   To try to replicate Derwas and Jones's findings. (Experiments 1 and 2.)

3.   To consider whether momentary DRO can be effective with subjects who have mild learning disabilities. (Experiments 3 and 4.)

4.   To consider whether momentary DRO can be effective with subjects who do not have learning disabilities. (Experiment 4.)

5.   To determine whether momentary DRO provides higher levels of reinforcement than whole-interval DRO (Experiment 1.)

EXPERIMENT 1

The relative effectiveness of momentary DRO and whole-interval DRO in the reduction of high-rate stereotyped behaviour

INTRODUCTION

The first study which showed momentary DRO to be at least as effective as whole-interval DRO was carried out by Derwas and Jones (1993). The present experiment was carried out in an attempt to determine whether these findings could be replicated, or if the contradictory findings of Repp et al. (1983) would be supported.

The aim of this experiment was to compare the effectiveness of momentary DRO and whole-interval DRO in the reduction of the high-rate stereotyped behaviour of two subjects. A multiple-baseline design was employed, and an additional stereotyped behaviour of each subject was monitored to assess any side-effects of the treatments carried out on the targeted behaviours. A record was also kept of the number of reinforcers subjects earned during each schedule, in order that any differences in reinforcer density could be assessed.

METHOD

Subjects

There were two subjects in this experiment: Matthew and Peter. Both subjects were residents at a small hostel for people with learning disabilities and inappropriate behaviour. Both subjects had severe learning disabilities and had only very limited self-help skills, requiring help with dressing, feeding and toileting. Matthew was 29 years old and Peter

was 26 years old. Neither subject had any speech, though they seemed able to understand simple commands.

Response definitions

Matthew's inappropriate behaviours included head-swaying, screaming, vocalising, clothes-removal, pinching (self), and regurgitating. Peter's behaviours included rocking, hand-gazing and pinching (others). It was decided to intervene with the highest rate stereotypies of each subject, and monitor their second highest rate behaviour as a control. Matthew's highest rate stereotypy was head-swaying, which was operationally defined as turning his head 180° from side to side. His second highest-rate stereotypy was his general vocalising (excluding screaming and shouting). This was operationally defined as a low-pitched continual moaning.

Peter's highest rate stereotypy was rocking, operationally defined as moving his whole torso vigorously back and forth whilst in a sitting position. His second highest stereotypy was hand-gazing, operationally defined as holding his hand in front of his face and starting at it for more than three seconds.

Ethical approval

For this experiment and all others in this thesis, appropriate ethical approval was obtained from the University of Wales, Bangor, School of Psychology Ethics Committee. Where appropriate, approval was obtained from the local Health Authority Ethics Committee. In addition, approval from the unit, institution or school, and parents, managers and care-workers was always obtained before studies were carried out.

Apparatus

During all phases of the experiment, a ten-second momentary time-sampling procedure was used to record the behaviours. The experimenter wore a set of headphones attached to a personal cassette player which contained a pre-recorded tape. The tape consisted of a voice

saying 'one', 'two', 'three' and so on at ten-second intervals, prompting the experimenter to take observations. The recording sheets used to note occurrences of the behaviours are reproduced in Appendix A, and a stop-watch was used to time the DRO intervals. A 'Y' plug and a second pair of head-phones allowed a second observer to make reliability checks at the same moments as the experimenter. An audio-cassette player and a tape of relaxation music were used to provide reinforcement.

Design

A multiple-baseline across subjects design was employed in this experiment. The baseline phase (A) for each subject was followed by one of two intervention phases, (B) whole-interval DRO and (C) momentary DRO. The order of presentation of the two interventions was counter-balanced across the two subjects in order to minimise order effects. Therefore the order of presentation for Matthew was A-B-C, whilst for Peter it was A-C-B. To measure response covariation effects, an additional stereotyped behaviour was monitored but not consequated for each of the subjects. Sessions were five minutes long. The experiment took place over a four-month period and the subjects were always seen together in the same room.

Reinforcer assessment

Care-staff were asked to complete Motivation Assessment Scales (Durand & Crimmins, 1988), to give an indication of the likely function of the stereotyped behaviours. Reinforcers were only chosen if they were identified by the MAS, direct observation *and* staff interview, as it is known that different methods of functional analysis disagree with each other (Crawford et al., 1992; Oliver & Head, 1993). Following this, potential reinforcers were subjected to a multiple-choice reinforcer assessment (e.g. Berg & Wacker, 1991) to select the most salient reinforcer for each subject. This assessment consisted of the following: four potential reinforcers were placed on a tray and this was presented to each subject several times over a two day period. The order of items on the tray was changed after each trial. At each trial the item which the subject chose (either by picking it

up or by reaching towards it) was noted, and at the end of the trials the item which had been chosen most often was used as the reinforcer during the intervention.

Three care-workers completed the Motivation Assessment Scale (MAS) for Matthew's head-swaying, and all agreed that the most likely maintaining variable was sensory reinforcement. Sensory reinforcement in the form of music was identified by care-staff as something Matthew responded to, so music was used together with the reinforcer most often chosen by him during the multiple-choice reinforcer assessment. Peter's key-worker completed the MAS for his rocking; this suggested sensory reinforcement was the most likely maintaining variable. However, sensory reinforcers were not identified either by observation or in staff questionnaires.

As a result of the various methods of reinforcer selection, the reinforcer chosen for Peter was orange, and the reinforcers chosen for Matthew were chocolate and music.

Inter-response times

Before baseline sessions began, a number of recordings of the length of time between occurrences of the two targeted behaviours were made. It was found that Matthew's head-swaying ceased, on average, for 14.55 seconds, whilst Peter's rocking ceased, on average, for 42.67 seconds. These times were used to set appropriate interval lengths for intervention: 12 seconds for Matthew and 30 seconds for Peter. These were slightly shorter than the mean inter-response times, in order to maximise the possibility of reinforcement, following procedures outlined by Jones (1991a).

Procedure

Baselines (A) ended in a staggered fashion once a level of stability had been achieved. When the first intervention (B) began for Matthew, baseline data continued to be collected for Peter for a further eight sessions, at which point intervention (C) was introduced for him. When sixteen sessions of intervention (B) had been completed with Matthew,

intervention (C) began. When sixteen sessions of intervention (C) had been completed with Peter, intervention (B) began. The experiment ended when both subjects had experienced sixteen sessions of each intervention.

(B) Whole-interval DRO.

During this intervention, reinforcers were presented if there had been no occurrence of the targeted behaviour during an interval. Reinforcer delivery was as follows: Matthew was offered a piece of chocolate, and once he had taken this or ignored it, ten seconds of music was played. Peter was offered a piece of orange. Timing of DRO intervals stopped while the music was playing, and while the subjects were eating.

If the target behaviour did occur during an interval, timing of that interval stopped, and the stop-watch was re-set to zero. When the target behaviour ceased, the stop-watch was restarted and a new interval began.

(C) Momentary DRO.

Reinforcers were presented if the target behaviour was not occurring at the moment when the interval ended, regardless of whether there had been any occurrences of the target behaviour during that interval. Delivery of reinforcers was carried out as for whole-interval DRO.

In this experiment, and in all others in this thesis (excepting Experiments 7, 9 and 10), reinforcement was not given in either DRO schedule if any undesirable behaviour was occurring at or just before the moment when reinforcement would otherwise have been given. This was to avoid the possibility of inadvertently reinforcing other inappropriate behaviours.

Inter-observer agreement -

For 33% of the sessions, observations were taken by a second observer to assess the reliability of the momentary time-sampling recording method. Inter-observer agreement was calculated using a point-by-point agreement ratio (Kazdin, 1982). This divides the number of points of agreement for each behaviour by the total number of points, as illustrated in the formula below.

Point-by-point agreement = (A / A + D) x 100

Where A = agreements for the point of observation

       D = disagreements for the point of observation

The results showed the following means and ranges for each of the four stereotyped behaviours measured: Matthew's head-swaying: mean = 96%, range = 87% to 100%; Matthew's vocalising: mean = 93%, range = 87% to 100%; Peter's rocking: mean = 99%, range = 93% to 100%; Peter's hand-gazing: mean = 88%, range = 77% to 100%.

RESULTS

Figure 1.1 shows the percentages of observations in which stereotyped behaviour occurred. For Matthew, head-swaying during baseline ranged from 7% to 88% and averaged 45%. In the whole-interval DRO intervention, head-swaying ranged from 0% to 97% and averaged 58%. In the final phase, momentary DRO, head-swaying ranged from 0% to 90% and averaged 34%.

For Peter, rocking during baseline ranged from 0% to 80% and averaged 39%. In the momentary DRO intervention rocking ranged from 0% to 47% and averaged 15%. In the final phase, whole-interval DRO, rocking ranged from 0% to 79% and averaged 17%.
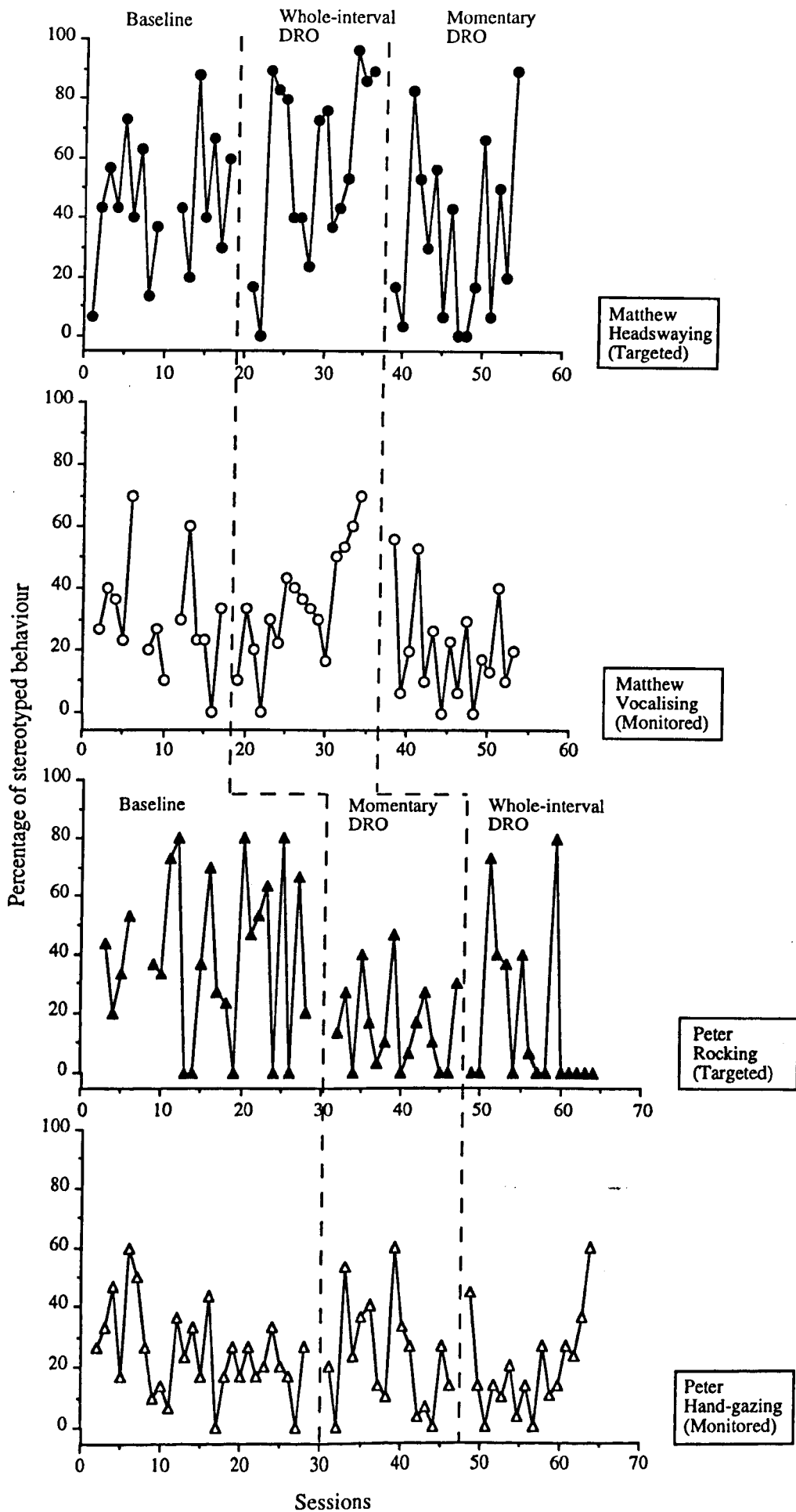
Figure 1.1 : The percentage of observations in which targeted and
monitored behaviours occurred.

Matthew's monitored stereotyped behaviour, vocalising, ranged from 0% to 70% during baseline and averaged 30%. During whole-interval DRO it also ranged from 0% to 70%, and averaged 34%. During momentary DRO it ranged from 0% to 57% and averaged 21%.

Peter's monitored stereotyped behaviour, hand-gazing, ranged from 0% to 60% in all three phases (baseline, momentary DRO and whole-interval DRO). It averaged 25% during baseline, 23% during momentary DRO, and 20% during whole-interval DRO.

Number of reinforcers

The number of reinforcers earned by subjects were recorded and are detailed in Table 1.1

Table 1.1 Number and range of reinforcers

	Whole-interval DRO		Momentary DRO	
	Mean	Range	Mean	Range
Matthew	2.25	0 - 8	7.31	2 - 11
Peter	4.94	0 - 7	5.31	3 - 7
Overall	*3.59*	*0 - 8*	*6.31*	*2 - 11*

Three unrelated t-tests were carried out on these data. The first tested the difference in the number of reinforcers across schedules for Matthew. This showed a highly significant difference ($t(30) = -5.381$, $p<0.001$ two-tailed), with considerably more reinforcers earned in momentary DRO than whole-interval DRO. The second t-test analysed the difference in the number of reinforcers for Peter. No significant difference was found here, although there were slightly more reinforcers earned in momentary DRO than in whole-interval DRO. The third t-test examined the difference in number of reinforcers between the two schedules overall, using the combined data for Matthew and Peter. A highly significant difference was found here ($t(62) = -4.195$, $p<0.001$ two -tailed), again with more reinforcers earned in momentary DRO than whole-interval DRO.

The results show that for Matthew, momentary DRO was more effective than whole-interval DRO in reducing the targeted stereotyped behaviour, head-swaying. Indeed, during whole-interval DRO head-swaying *increased* relative to baseline. As soon as the momentary DRO schedule was introduced, the frequency of head-swaying fell to levels below baseline.

It is not clear why whole-interval DRO caused Matthew's target behaviour to increase. It is possible that the low levels of contingent reinforcement meant that Matthew had limited access to stimulation other than his stereotypy. This is supported by the decrease in target behaviour seen under momentary DRO, during which schedule he earned far higher levels of reinforcement than in whole-interval DRO. The occasional tendency for whole-interval DRO to increase behaviour has been noted before (e.g. Friman et al., 1986), and is considered in more detail in Chapter 3.

For Peter, both schedules decreased his targeted stereotyped behaviour (rocking) relative to baseline. Although there was greater variability in rocking during whole-interval DRO, most of the highest sessions were at the start of the intervention. It may be therefore that this variability was caused by the change between the two interventions rather than the intervention itself. Certainly during the last few sessions of whole-interval DRO, rocking was totally absent.

During the momentary DRO intervention, Matthew's monitored stereotyped behaviour, vocalising, was also substantially reduced, compared to levels during both baseline and whole-interval DRO. The pattern on the graph (Figure 1.1) suggests that Matthew's two highest-rate stereotypies tended to occur concurrently: as head-swaying increased in whole-interval DRO, so did vocalising, whilst both fell during momentary DRO. Informal observation supports this; Matthew tended to combine head-swaying and vocalising.

There appeared to be little-or no effect of the interventions on Peter's monitored stereotyped behaviour (hand-gazing). Whilst it appears from the graph that hand-gazing was occurring at higher levels during momentary DRO than baseline, the percentage data show that the range during both phases was identical, and the mean percentage of stereotypy was in fact very slightly lower during momentary DRO than baseline. The mean amount of hand-gazing was lowest in the whole-interval DRO phase, although the range was identical to that in the other two phases. It seems that there was a certain amount of covariation between Peter's two stereotyped behaviours: in momentary DRO, where rocking was low, hand-gazing was comparatively high, whilst in whole-interval DRO, where rocking was reduced to zero, hand-gazing started to increase. As with Matthew, informal observation confirms this pattern. Peter tended to stop rocking in order to hand-gaze, which meant that the two behaviours did not often occur simultaneously.

The fact that differences were seen in the monitored behaviour patterns of Matthew and Peter suggests that studies which have not taken similar measures may have been unable to gauge the full effects of the treatments they used. This is particularly relevant when other behaviours are injurious to the subject or others. A treatment which reduces stereotyped behaviour but produces concomitant increases in self-injury is of little value. Unless measures of other behaviours are taken, this cannot be fully assessed.

Overall, considerably more reinforcers were delivered during momentary DRO than whole-interval DRO. It is interesting to note, however, that the individual results for Peter showed little difference between schedules regarding the number of reinforcers. Peter earned relatively high levels of reinforcement during whole-interval DRO. This may have been partly due to the fact that the DRO interval for Peter - 30 seconds - was rather less than his average inter-response time - 42.67 seconds. This may have allowed his behaviour to contact the contingencies at a relatively high rate.

However, the results for Matthew and the results overall show a large difference between schedules in reinforcement levels. There were also differences in the ranges of reinforcement. In whole-interval DRO, Matthew earned zero reinforcers in 50% of sessions, and Peter earned zero reinforcers in 12.5% of sessions. In momentary DRO, however, neither subject earned zero reinforcers. Matthew never earned less than two reinforcers and in one session earned eleven. In 50% of sessions in momentary DRO Matthew earned eight or more reinforcers (eight was the highest amount he earned in whole-interval DRO). Peter never earned less than three reinforcers in momentary DRO, although the highest amount he earned was the same as for whole-interval DRO (seven). It may be worth noting that because of the different interval lengths used for each subject, they had a different number of potential opportunities each session to earn reinforcers. For each five minute session, Matthew had the opportunity to earn 25 reinforcers (his DRO interval length being 12 seconds), while Peter had the opportunity to earn just 10 (as his DRO interval length was 30 seconds). However, this discrepancy seems to have had little impact on the actual rates of reinforcement: indeed, during whole-interval DRO Peter earned far higher levels of reinforcement than Matthew. During momentary DRO, Matthew earned slightly more.

Despite the individual differences, the results overall show that momentary DRO produced greater levels of reinforcement than whole-interval DRO. This raises the question of whether the effectiveness of momentary DRO was due to an increased opportunity for reinforcement. Peter earned similar amounts of reinforcement in both schedules, and there was little difference in the rate of his behaviour between the two schedules. Matthew, on the other hand, earned far greater amounts of reinforcement in momentary DRO and for him, this schedule was more effective than whole-interval DRO. The possible effects of the higher levels of reinforcement in momentary DRO are considered in later experiments in this thesis.

Overall, then, the results of this experiment support those of Derwas and Jones (1993). They show that momentary DRO can be as effective, and sometimes more effective, than whole-interval DRO in reducing the high-rate stereotyped behaviour of some individuals.

EXPERIMENT 2

The relative effectiveness of momentary DRO and whole-interval DRO in the reduction of high-rate stereotyped behaviour: An alternating treatments design

INTRODUCTION

Repp et al. (1983) carried out two studies to compare the relative effectiveness of momentary DRO and whole-interval DRO. Their first study, like the previous experiment in this chapter, employed a multiple-baseline design. Their second study employed a multiple schedule design as an alternative way of comparing the two schedules. This study found that whole-interval DRO was more effective than momentary DRO in reducing the disruptive behaviour of an individual with moderate learning disabilities. Following their methodology, the present experiment employed a similar design in order to compare the schedules in the reduction of high rate stereotyped behaviour.

The design used in the present experiment made four changes to the design used by Repp et al. (1983). First, Repp et al. attached a different coloured piece of paper to the subject's desk to accompany each of the three conditions, so that the start of a new session was effectively announced by the new piece of paper. This is one of the characteristics of a multiple-schedule design: each condition is associated with a different setting or some other factor which, it is believed, will encourage discrimination between conditions. Although associating a different colour with a different schedule might enhance discrimination, it does add a potentially confounding variable to the intervention. It was felt to be simpler and allow less scope for ambiguous results if schedules were alternated randomly and contained no differences other than those inherent in the schedules themselves. This study therefore employed an alternating treatments design, which does not require different

treatments to be associated with different factors. In other respects it is identical to a multiple-schedule design.

Second, Repp et al. randomly alternated baseline, momentary DRO and whole-interval DRO sessions. It was decided for the present study to take baseline measures prior to intervention, for the following reason. If baseline is one of the 'treatments' that is measured concurrently with other treatments, a problem arises if there is no separation on the graph between baseline and intervention. In such a case it would be impossible to determine whether this was due to a contamination effect between baseline and treatment, or simply whether treatment had no effect on behaviour. An initial baseline was taken in the present study, therefore, as it was felt that this would allow stronger assumptions to be made.

The two other differences between the present design and that of Repp et al. were (i) that other inappropriate behaviours were recorded, but not consequated, to assess any response covariation; and (ii) DRO interval lengths were based on the naturally occurring frequency of the target behaviour rather than the standard five minute interval used by Repp et al.

METHOD

Subject

Paddy was 28 years old with severe learning disabilities (his mental age was approximately two years, as measured by the Vineland Social Maturity Scale). He was a resident at the same home as Matthew and Peter. Paddy suffered from epilepsy and was on medication to control this. Paddy possessed some self-help skills: he could feed and toilet himself, though he required some help in dressing. He could speak a few words, and seemed able to understand a number of instructions.

Response definition

Paddy presented a variety of inappropriate behaviours, including hand-biting, rocking, face-slapping, continuously rattling objects and shouting. As with Experiment 1, it was decided to intervene with Paddy's highest rate stereotypy which was rocking, operationally defined as moving his whole torso rapidly backwards and forwards whilst in a sitting position. Hand-biting and face-smacking were chosen as behaviours to be monitored but not consequated; as these were self-injurious behaviours it was particularly important to ascertain whether they were effected by the intervention. These three behaviours were agreed upon after consultation with care-staff.

Apparatus

A personal cassette-player and headphones, pre-recorded audio-tape, recording sheets and stop-watch were used in this experiment, as described in Apparatus section for Experiment 1.

Design

An alternating treatments design was employed in this experiment. A baseline phase was followed by an intervention phase where two treatments - momentary DRO and whole-interval DRO - were randomly alternated (following procedures outlined by Kazdin, 1982). Each five minute session was immediately followed by the next, according to a pre-determined list of treatments. To measure response covariation effects, two self-injurious behaviours were monitored but not consequated. A final phase was carried out in which the most effective treatment was continued alone for six sessions.

Reinforcer

Consultation with care-staff showed that Paddy was highly motivated by food, and on their recommendation five foods were selected as possible reinforcers. A multiple-choice reinforcer assessment was then carried out, where the five foods were presented to Paddy over ten trials, as described in Experiment 1. Following this procedure, chocolate Smarties

were found to be chosen more than any other items and these were therefore used during the experiment.

Inter-response time

Before baseline sessions began, the length of time between occurrences of the targeted behaviours was recorded over several sessions. It was found that Paddy's rocking ceased on average for 30 seconds. This was used to set a DRO interval of 25 seconds.

Procedure

Baseline

Baseline measures of the target behaviour (rocking) were taken until there was stability. This occurred after 14 sessions.

Whole-interval DRO

During this intervention Paddy was given a reinforcer if he had not rocked at all during the interval of 25 seconds. Timing of the DRO interval stopped while he ate the reinforcer.

If he did rock during an interval, timing of that interval ceased and the stop-watch was reset. A new interval began when the rocking ceased. Seventeen sessions of this intervention were carried out.

Momentary DRO

During this intervention a reinforcer was given if Paddy was not rocking at the moment between two intervals, regardless of whether the behaviour had been displayed during the interval. Timing of the DRO interval stopped while he ate the reinforcer. If he was rocking at that moment, reinforcement was not given, the stop-watch was reset and a new interval began. Seventeen sessions of this intervention were carried out.

Continuation phase

The most successful intervention was continued in this phase for six sessions.
Throughout baseline and intervention phases, recordings were made of the occurrences of
the two monitored behaviours.

Inter-observer agreement

For 9% of the sessions, observations were taken by a second observer. Inter-observer
agreement was calculated using a point-by-point agreement ratio (Kazdin, 1982) as outlined
in Experiment 1. The results showed the following means and ranges for each of the three
behaviours measured: rocking: mean = 84%, range = 80% to 97%; hand-biting: mean =
97%, range = 93% to 100%; face-smacking: mean = 87%, range = 77% to 97%.

RESULTS

Figure 2.1 shows the percentage of observations in which inappropriate behaviour
occurred. Paddy's targeted stereotypy, rocking, ranged during baseline from 7% to 60%
and averaged 31%. During whole-interval DRO, rocking ranged from 7% to 30% and
averaged 16%. During momentary DRO, rocking ranged from 0% to 27% and averaged
12%. These means show that both schedules reduced the target behaviour and that there
was little difference between them. To confirm whether there was a statistical difference,
an approximation to the randomization test (Kazdin, 1982) was carried out. This took the
form of a t-test. It was possible to use this test because of the random alteration of
treatments (for an outline of the rationale behind this test, see-Appendix B). The t-test
showed that there was not a significant difference between momentary DRO and whole-
interval DRO (t(32) = 1.239, p>0.1).
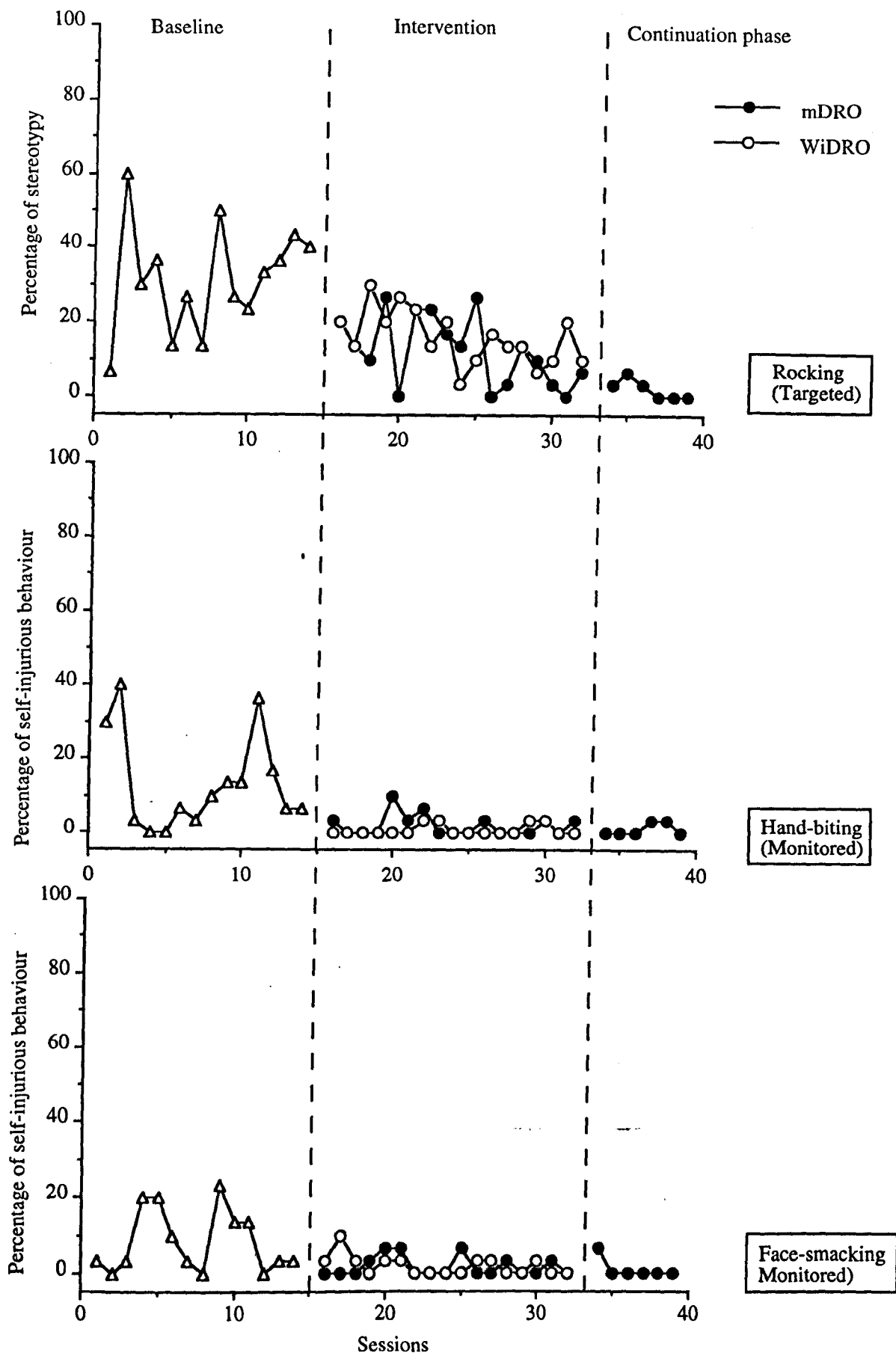
Figure 2.1 : The percentage of observations in which Paddy's behaviours occurred.

N.B.   In sessions where identical rates of behaviour occurred in both schedules, the white circle representing whole-interval DRO is printed over the top of the black circle representing momentary DRO.

The first of Paddy's monitored behaviours, hand-biting, ranged from 0% to 40% during baseline and averaged 13%. During whole-interval DRO it ranged from 0% to 3%, and averaged 0.8%. During momentary DRO it ranged from 0% to 10% and averaged 2%. A t-test carried out on these data showed that there was not a significant difference between schedules ($t(32) = -1.494$, $p > 0.1$).

Paddy's second monitored behaviour, face-smacking, ranged from 0% to 23% and averaged 8%. During whole-interval DRO it ranged from 0% to 10% and averaged 2%. During momentary DRO it ranged from 0% to 7% and averaged 2%. A t-test carried out on these data showed that there was not a significant difference between schedules ($t(32) = 0.216$, $p > 0.1$).

Momentary DRO was continued in the final phase for six sessions. In this phase, rocking ranged from 0% to 7% and averaged 2%. Hand-biting ranged from 0% to 3% and averaged 1%. Face-smacking ranged from 0% to 7% and averaged 1%.

DISCUSSION

The results show that both whole-interval DRO and momentary DRO were effective in reducing Paddy's targeted behaviour, rocking. Although there was not a statistically significant difference between the schedules, momentary DRO had a lower overall mean than whole-interval DRO, and also contained three sessions of zero stereotypy (whole-interval DRO did not contain any). Momentary DRO was therefore continued alone for six sessions, during which time it maintained a low rate of rocking.

Both interventions also produced sizeable reductions in the two monitored behaviours. Whole-interval DRO reduced hand-biting to zero in 76% of sessions, whilst momentary DRO produced a more variable effect. However, the overall difference between the schedules was negligible. Whole-interval DRO reduced face-smacking to zero in 53% of

sessions, whilst momentary DRO reduced it to zero in 35% of sessions. However, momentary DRO produced a slightly greater reduction overall. T-tests showed that there were no significant differences between the two schedules. During the continuation phase, momentary DRO maintained suppression of the two monitored behaviours for a further six sessions.

The effect of the chocolate reinforcers diminished for Paddy over each day, presumably as the result of satiation. Often, when he was offered a Smartie he would not eat it; instead he would put it in a tin and shake it. However, not eating the chocolate did not have any noticeable effect on the reduction of Paddy's inappropriate behaviours. This suggests strongly that the attention and close contact provided by the experimenter was a powerful reinforcer for him which competed successfully with the reinforcing qualities of his stereotypy. Although the unit where the experiment was carried out was well-staffed, one-to-one continual attention for one client was rarely possible. The continual attention provided during intervention sessions, therefore, was something of a novelty to Paddy. Informal observation confirmed that he greatly enjoyed, and often sought, close contact with staff. This attention, then, appeared to be the reason for the reduction in Paddy's inappropriate behaviours. The slightly greater reduction in momentary DRO may have been due to the greater density of reinforcement that this schedule tends to provide.

Studies which have carried out component analyses of schedules have found similar results. For example, Jones and Baker (1989) showed non-contingent attention to be the main reductive component of a DRI schedule that they had used with three subjects. In the present experiment, the two monitored behaviours were simultaneously suppressed when rocking was suppressed. These results suggest that attention was the maintaining variable for all Paddy's inappropriate behaviours, and that the programmed reinforcer was an arbitrary reinforcer. The role of attention in the reduction of inappropriate behaviour is complex and has implications for the validity of functional analysis. This is considered in more detail in the next chapter.

The results of this study show that momentary DRO was slightly more effective than whole-interval DRO in reducing high-rate stereotyped behaviour. However, both schedules were highly effective. This study also shows that both schedules can have an equally suppressive effect on inappropriate behaviour not directly targeted by the intervention.

EXPERIMENT 3

The relative effectiveness of momentary DRO and whole-interval DRO in the reduction of lower-rate disruptive behaviour

INTRODUCTION

The previous two experiments compared the relative effectiveness of momentary DRO and whole-interval DRO on high-rate stereotyped behaviours. The present experiment was designed to assess the schedules' effectiveness on lower-rate inappropriate behaviour. The behaviours chosen for intervention were behaviours which are categorised as 'classroom disruption': inappropriate physical contact with peers and lack of attention.

This experiment was a development of the first study of Repp et al. (1983), who compared the two schedules' effectiveness in reducing classroom disruption in children with mild-to-moderate learning disabilities. There were three differences between their study and the present experiment.

Firstly, as outlined in Chapter 1, it may be that Repp et al. found little effect of momentary DRO because they made such a clear and specific connection between a particular signal (a raised hand) and the subject earning a reinforcer if not disrupting at that signal. This informed their subjects clearly that they only had to be 'good' at that signal in order to earn a reinforcer. The present experiment made use of less specific connections. Subjects were given clear and complete instructions, but their attention was not drawn to the relationship between a particular moment and a reinforcer. This allowed a clearer assessment of the effects of momentary DRO on children who were capable of using rules.

Secondly, this experiment made use of DRO intervals based on the rate of the behaviour, while Repp et al. used a standard five-minute interval for all their subjects. Thirdly, this experiment used a resetting whole-interval DRO schedule, while Repp et al. used a non-resetting interval.

Repp et al. (1983) worked with children who had mild to moderate learning disabilities, and they found momentary DRO to be less effective than whole-interval DRO. They proposed momentary DRO would be even less effective with more capable subjects. The subjects in this experiment had mild learning disabilities, to allow an examination of this suggestion.

METHOD

Subjects

There were two subjects in this experiment: Timothy and Paul. Both subjects had mild learning disabilities and were part-time students at an Emotional and Behavioural Difficulties unit. Both subjects attended mainstream schools three days each week. Timothy was seven years old and Paul was five years old.

Response definitions

Timothy was inclined to be disruptive and occasionally was aggressive. His disruptive behaviour included inappropriate contact with his peers during classroom activities, such as tapping, touching and kicking. He also swore and made faces. Paul exhibited a general lack of concentration. He displayed most of the 'off-task' behaviour described by Repp et al. (1983), including off-task physical behaviour, such as pencil tapping; off-task attention, whereby he continually looked away from his work; off-task verbal behaviour, where he talked about topics unrelated to the task; and off-task interruptions where he would interrupt someone to talk about a topic unrelated to the task. He also fidgeted a great deal and did not sit properly in his chair.

After discussion with teaching staff, it was decided to intervene with Timothy's inappropriate physical contact, and Paul's off-task attention. Staff suggested that it was Paul's lack of attention to the task in hand which triggered all the other disruptive behaviours. Inappropriate physical contact was operationally defined as touching anyone or distracting them unless touching them or talking to them was relevant to the task. Off-task attention was operationally defined as looking away from the task around the classroom and/or fidgeting.

Apparatus

Recording sheets were designed which allowed a line to be drawn through slots representing ten second periods (see Appendix A). In this way the duration of a behaviour could be recorded. Two stop-watches were used, one to record the DRO interval length and the other to record the duration of the inappropriate behaviour.

Design

A multiple-baseline across subjects design was employed in this experiment. The baseline phase (A) for each subject was followed by one of two intervention phases, (B) momentary DRO and (C) whole-interval DRO. The order of presentation of the two interventions was counter-balanced across the two subjects. Therefore the order of presentation for Timothy was A-B-C, whilst for Paul it was A-C-B. Sessions were five minutes long. The experiment took place in a classroom during normal class activities with six other children present.

Reinforcers

Teaching staff completed the Motivation Assessment Scale for each subject. These suggested that the most likely maintaining variable for Timothy's disruption was tangible reinforcement: he tended towards disruption particularly if a favourite toy or activity was taken away from him. However, attention was also rated very highly. The most likely maintaining variable for Paul was sensory stimulation; when left to himself for long periods

he would tend to fidget and look round the room more. However, attention was also rated highly for him.

Staff also completed a 'reinforcer suggestions' questionnaire (adapted from Martin & Pear, 1992). Overall, they recommended that both subjects would respond well to earning stickers - a gold star for Timothy and stickers of the cartoon character 'Mr Blobby' for Paul. This method of reinforcement was convenient as it could be used fairly discreetly in a classroom setting. A star system for merit was already used by the class, and Paul had previously been given 'Mr Blobby' stickers when he first joined the unit in order to encourage him to sit still during group work. It was decided to combine the stickers with praise, as this was a common aspect of classroom practice and staff felt that the children were responsive to it. This took into account both the tangible and attention factors which the MAS results had suggested as motivators.

Inter-response times
Baseline data were used to calculate the mean length of time between disruptive responses. The mean inter-response time for Timothy was found to be 1 minute 57 seconds, which was rounded up to two minutes as it was felt this would be a more workable interval. The mean inter-response time for Paul was found to be 1 minute and 20 seconds.

Procedure
Prior to baseline, the experimenter spent some time in the classroom so that the children could become accustomed to her presence. Once staff felt that the childrens' behaviour had resumed to normal, baseline commenced. Baselines (A) ended in a staggered fashion once a level of stability had been achieved. When the first intervention (B) began for Timothy, baseline data continued to be collected for Paul for a further two sessions, at which point intervention (C) was introduced for him. When twelve sessions of intervention (B) had been completed with Timothy, intervention (C) began. When twelve sessions of

intervention (C) had been completed with Paul, intervention (B) began. The experiment ended when both subjects had experienced twelve sessions of each intervention.

(B) Momentary DRO

Before this intervention began, subjects were given verbal instructions. Timothy was told, "I'm going to look at you every two minutes and if you're being good, not touching anyone or distracting anyone I'll give you a gold star." Paul, who was younger, was given a clock to put on his desk, and the experimenter moved the hands round to show him five minutes, and explained that he could earn three stickers during that time if he was concentrating when the experimenter looked at him. Both subjects were asked to repeat back to the experimenter what was required of them.

Subjects earned a reinforcer if the target behaviour was not occurring at the moment when the interval ended, regardless of whether there had been any occurrences of the target behaviour during that interval. The experimenter would look up and deliver verbal praise such as, "Good boy... you're behaving really well". Occasionally the verbal praise for Paul was more explicit, to remind him what he had to do to earn a sticker. This praise might be something like, "Well done, you behaved really well then. You were looking at your work and not around the classroom. You also weren't fidgeting but were concentrating on what you were supposed to be doing, so that's another blob you've got to stick on Mr Blobby. Let's see if you can get some more."

The verbal praise was a sign that the subject had earned a sticker. Timing of the DRO interval was suspended until the praise had been completed. Stickers were given to the subjects at the end of the five minute session. Timothy was allowed to stick his stars into a book at break-time, while Paul was allowed to place the stickers he earned onto a 'Mr Blobby' wall-chart. There were small and large stickers and Paul could choose which ones he wanted each time.

(C) Whole-interval DRO

Before this intervention, Timothy was told, "Now to get your stars you have to be good and not distract or touch anyone for two minutes. If you do distract or touch someone then I'll begin timing the two minutes again until you can be good for all that time."

Paul was told, "I'm going to be watching you to see if you can concentrate on your work and what you're asked to do. If you can concentrate and not play with pens, or talk, or look up and fidget for a whole minute and twenty seconds, I'll say 'good boy' and I'll put a blob aside for you to stick on Mr Blobby later. If you can't concentrate then your minute and twenty seconds will start again. Until you can be good for all of that time you won't get a blob for Mr Blobby. So just be good and concentrate on your work for as long as you can."

Both subjects were asked to repeat back to the experimenter what was required of them.

During this phase, subjects had to refrain from the target behaviour for the entire interval in order to earn a reinforcer. If the target behaviour occurred during an interval, timing of that interval stopped, and the stop-watch was re-set to zero. When the target behaviour ceased, the stop-watch was restarted and a new interval began.

If subjects refrained from the target behaviour for the entire interval, the experimenter praised them and put aside a sticker for them to have at the end of the session, as for momentary DRO.

RESULTS

Figure 3.1 shows the percentages of the duration of the disruptive behaviour in each session. For Timothy, inappropriate physical contact during baseline ranged from 0% to 30% and averaged 8%. In the momentary DRO intervention, inappropriate contact
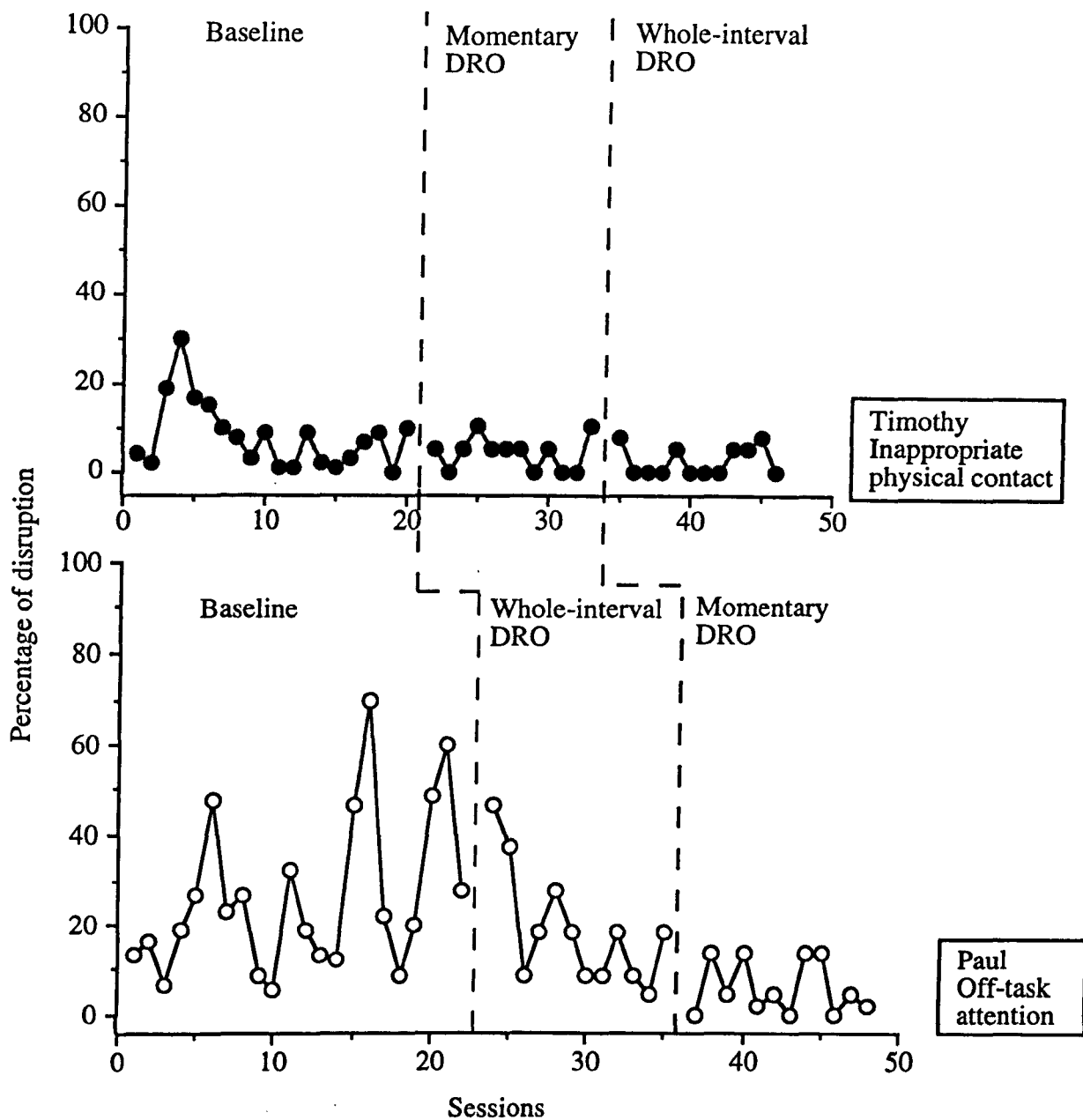
Figure 3.1 : The percentages of the duration of the targeted behaviours.

ranged from 0% to 10% and averaged 4%. In the whole-interval DRO intervention,

inappropriate contact ranged from 0% to 8% and averaged 3%.

For Paul, off-task attention during baseline ranged from 6% to 70% and averaged 26%.

In the whole-interval DRO intervention off-task attention ranged from 4% to 47% and

averaged 19%. In the momentary DRO intervention off-task attention ranged from 0% to 14% and averaged 6%.

DISCUSSION

The mean data suggest that both interventions reduced Timothy's inappropriate physical contact. However, Figure 3.1 shows that intervention began when baseline was at low levels, and both interventions appear simply to be continuations of baseline. The mean data for Paul suggest that momentary DRO produced more substantial reductions than whole-interval DRO. However, the graph depicts a clear downward trend across intervention phases, which means that it cannot be concluded from these data that momentary DRO was more effective than whole-interval DRO.

It seems then that regardless of schedule, the disruptive behaviour of both subjects decreased during the course of the study. Timothy's graph shows that the relatively high level of disruption during baseline gradually declined, so that momentary DRO and whole-interval DRO are both simply continuations of these low levels. Indeed, it cannot be stated confidently that either or both interventions were successful in reducing Timothy's disruption, as the latter half of his baseline was so low. Any effect of the interventions would be difficult to detect, as the behaviour was already fairly near floor level. At the very least, though, both interventions did seem to have a stabilising effect on Timothy's behaviour, as the range of disruption during both interventions was considerably lower than during baseline.

Similarly, with Paul's graph there is a gradual decrease across the interventions so that behaviour during momentary DRO appears to be an extension of the pattern seen during whole-interval DRO.

For Paul, the highest occurrences of disruption during baseline appeared to coincide with table work (such as writing or mathematics), where there were materials available to fidget with, such as pens and chalk. It is possible that the off-task attention was due in part to boredom or frustration at being unable to do the task. It seemed towards the end of the study that Paul had made a connection between 'being good' and getting stickers; he told the experimenter how 'good' he had been during the week at his mainstream school. Towards the end of the study Paul left the unit and was reintegrated into his mainstream school full-time, because his behaviour had improved so considerably. It appears that the DRO intervention contributed to this improvement.

Repp et al. (1983) hypothesised that momentary DRO would not work with capable subjects who would be able to easily discriminate between the two schedules. This experiment does not support this hypothesis. Although the instructions outlined the difference between the schedules, momentary DRO was as effective as whole-interval DRO. This may have been because the 'moment' was not signalled as Repp et al. had done; and thus there was less emphasis on the importance of a single moment. This suggests that it is possible to give clear instructions regarding the nature of momentary DRO without eliminating its effectiveness. It is possible, though, that the subjects did not fully understand the instructions. At the very least, this experiment questions the suggestion that momentary DRO is unsuitable for people who have only mild learning disabilities.

The results of this study, then, did not support the findings of Repp et al. They reported that momentary DRO was not strong enough to produce a reduction in disruptive behaviour, although it could maintain a reduction begun by whole-interval DRO. In this study, momentary DRO was as effective as whole-interval DRO in reducing disruptive behaviour.

Because Timothy's disruptive behaviour had decreased during baseline, and there was a downward trend in Paul's data, the effect of the schedules was not clear. It was therefore decided to carry out a second experiment to compare the two schedules in the reduction of lower rate inappropriate behaviour. This experiment is described overleaf.

EXPERIMENT 4

The relative effectiveness of momentary DRO and whole-interval DRO in the reduction of lower-rate disruptive behaviour: A development of Experiment 3

INTRODUCTION

This experiment developed further the issues raised in Experiment 3, which found that momentary DRO was as effective than whole-interval DRO in reducing disruptive behaviour. However, there were trends in the data which made a comparison between the schedules impossible. Therefore, it was decided to carry out a further experiment to compare the effectiveness of the two schedules. The present experiment worked with two subjects, one of whom had mild learning disabilities, the other of whom did not have learning disabilities. This was to further test the contention of Repp et al. that momentary DRO would be ineffective for capable subjects.

The present study kept a record of one other disruptive behaviour, as well as engagement on task, to allow examination of any response covariation. In addition, follow-up measures were taken three months after the end of the intervention to assess the generalisation of the interventions over time. A non-resetting whole-interval DRO schedule was used, as this was used by Repp et al., and it was considered to be more practical than a resetting schedule when recording the duration of behaviour.

The results of Experiment 3 suggested that it may have been the signal that Repp et al. used to indicate the end of an interval that contributed to the poor performance of momentary DRO. Therefore it was decided to use a signal in this experiment to determine if this would cause momentary DRO to be less effective than whole-interval DRO.

METHOD

Subjects

There were two subjects in this experiment: Perry and Mark. Both were seven years old, and they attended a mainstream junior school. Perry had mild learning disabilities; Mark did not have learning disabilities.

Response definitions

Perry constantly disrupted the class and sought attention continually from the teacher. Mark's parents and teachers described him as a 'day-dreamer' who spent large parts of the day staring into space, and tended to be thinking about something other than the class activities. It was decided to intervene with the most frequently occurring behaviours. For Perry this was his verbal disruption, which was operationally defined as:

(i)   Verbally interrupting another student's concentration on the classroom task with an unrelated task or comment

(ii)   Talking when the teacher does not require it

(iii)   Talking about a task unrelated to the class work

Mark's 'day-dreaming' was operationally defined as stopping the task and looking away to observe another student or to stare out of the window.

Two behaviours were selected to be monitored in order to assess response covariation. These were the second most frequently occurring disruptive behaviour, and on-task behaviour. The behaviour selected as Perry's monitored disruptive behaviour was getting out of his seat, and for Mark it was disrupting others. The definition of on-task behaviour was taken from Groden (1989): co-operative, pro-social behaviour such as working on assignment, sitting attentively and talking appropriately (positively) with teachers and peers.

Apparatus

Recording sheets were used to record the duration of the targeted, monitored and on-task behaviours (see Appendix A). A stop-watch was used to record the DRO interval length and the duration of the behaviours.

Design

A multiple-baseline across subjects design was employed in this experiment. As in Experiment 3, the baseline phase (A) for each subject was followed by one of two intervention phases, (B) whole-interval DRO and (C) momentary DRO. The order of presentation of the two interventions was counter-balanced across the two subjects. Therefore the order of presentation for Perry was A-B-C, whilst for Mark it was A-C-B. To measure response covariation effects, an additional disruptive behaviour and on-task behaviour were monitored but not consequated for each of the subjects. Sessions were fifteen minutes long. The experiment took place over four weeks in a classroom during normal class activities with 27 other children present. Three months later, a two-day follow-up study was carried out.

Reinforcer assessment

Possible reinforcers were chosen after discussion with staff and parents, and these were then tested experimentally over a series of trials in a procedure adapted from Repp et al. (1983). These trials tested the effectiveness of the stimuli in increasing the time given to a particular class activity (a drawing and colouring task). The item which produced the greatest increase in this activity for both subjects was multi-coloured sticky stars, and so these were used as reinforcers in this experiment.

Inter-response times

During baseline, the mean length of time between responses was calculated. Perry's mean inter-response time was 1 minute 3 seconds, and Mark's was 2 minutes 5 seconds. These

were rounded down slightly, so that the DRO interval for Perry was set at 1 minute, while for Mark it was 2 minutes.

Procedure

Prior to baseline, the experimenter spent some time observing the subjects in the classroom so that the children could become accustomed to his presence.

As in Experiment 3, baselines (A) ended in a staggered fashion contingent upon the achievement of baseline stability. When the first intervention (B) began for Perry, baseline data continued to be collected for Mark for a further four sessions, at which point intervention (C) was introduced for him. When nine sessions of intervention (B) had been completed with Perry, intervention (C) began. When ten sessions of intervention (C) had been completed with Mark, intervention (B) began.

(B) Whole-interval DRO.

Verbal instructions were given to the subjects prior to the start of this intervention. Perry was told, "If you do not disturb your class-mates and stop them from working throughout the whole one minute you will be allowed to stick a star on your chart." Mark was told, "If you are not staring into space for the whole of the next two minutes you will be allowed to stick a star on your chart." The end of the interval was signalled to Mark by a tap on the table. However, Perry was not given a signal as it was quickly noted that he became aggressive if he found that an interval had ended and he had not earned a reinforcer.

During this intervention, reinforcers were presented if there had been no occurrence of the targeted behaviour during an interval. Timing of the DRO interval was suspended while the subjects stuck their stars on the chart. If the target behaviour occurred during an interval the stop-watch was not re-set. However, no reinforcer was provided at that end of that interval, and timing then started again for the next interval.

(C) Momentary DRO

Verbal instructions were given to the subjects at the start of this intervention. Perry was told, "If you are not disturbing your class-mates and stopping them from working at the moment one minute ends you will be allowed to stick a star on your chart." Mark was told, "If you are not staring into space when I tap the table you will be allowed to stick a star on your chart."

During this intervention, reinforcers were presented if the target behaviour was not occurring at the moment an interval ended. Timing of the DRO interval was suspended while the subjects stuck their stars on the chart.

Follow-up

Three months after the study ended, data were collected as in the baseline phase, for four sessions.

Inter-observer agreement

The school, wishing to minimise disruption to classroom routine, would not allow another observer to join the experimenter in order to assess reliability. Therefore for a small percentage of the observations (just over 6%) a Welfare Worker, who visited the class infrequently, acted as a second observer.

Inter-observer agreement was calculated using an adaptation of Kazdin's (1982) point-by-point agreement ratio, as outlined in Experiment 1. A strict definition of agreement was drawn up, defined as (i) agreement that a behaviour had occurred and (ii) agreement of the length of behaviour to the nearest ten seconds. Both these criteria had to be achieved for agreement to be recorded. The results showed the following means and ranges for the observed behaviours: Perry's verbal disruption: mean = 79%, range = 72% to 85%; Perry's on-task behaviour: mean = 81%, range = 75% to 86%; Mark's day-dreaming: mean = 45%, range = 33% to 57%; Mark's disruption : mean = 88%, range = 75% to

100%; Mark's on-task behaviour : mean = 73%, range = 71% to 76%. Perry's out of seat behaviour did not occur while reliability checks were being made.

RESULTS

Figure 4.1 shows the six behaviours which were recorded in this experiment. Although necessary to show the staggered baselines, the size of the figure does not allow easy visual analysis. Therefore the two subjects' graphs have been separated and are presented below (Figures 4.2 and 4.3).

Perry's target behaviour, disruption, averaged 25% in baseline, and ranged from 7% to 40%. During whole-interval DRO, disruption averaged at 29% and ranged from 0% to 85%. During momentary DRO, disruption averaged at 39% and ranged from 15% to 72%.

Perry's monitored behaviour, 'out of seat', averaged at 8% in baseline and ranged from 0% to 20%. During whole-interval DRO, 'out of seat' averaged at 7.5% and ranged from 0% to 46%. During momentary DRO, 'out of seat' averaged at 2% and ranged from 0% to 11%. Perry's on-task behaviour averaged at 54% in baseline and ranged from 32% to 93%. During whole-interval DRO, on-task behaviour averaged at 63% and ranged from 10% to 100%. During momentary DRO, on-task behaviour averaged at 59% and ranged from 28% to 85%.

During the follow-up sessions Perry's disruption averaged at 70% and ranged from 59% to 82%. 'Out of seat' averaged at 3% and ranged from 0% to 9%. On-task behaviour averaged at 22% and ranged from 12% to 30%.
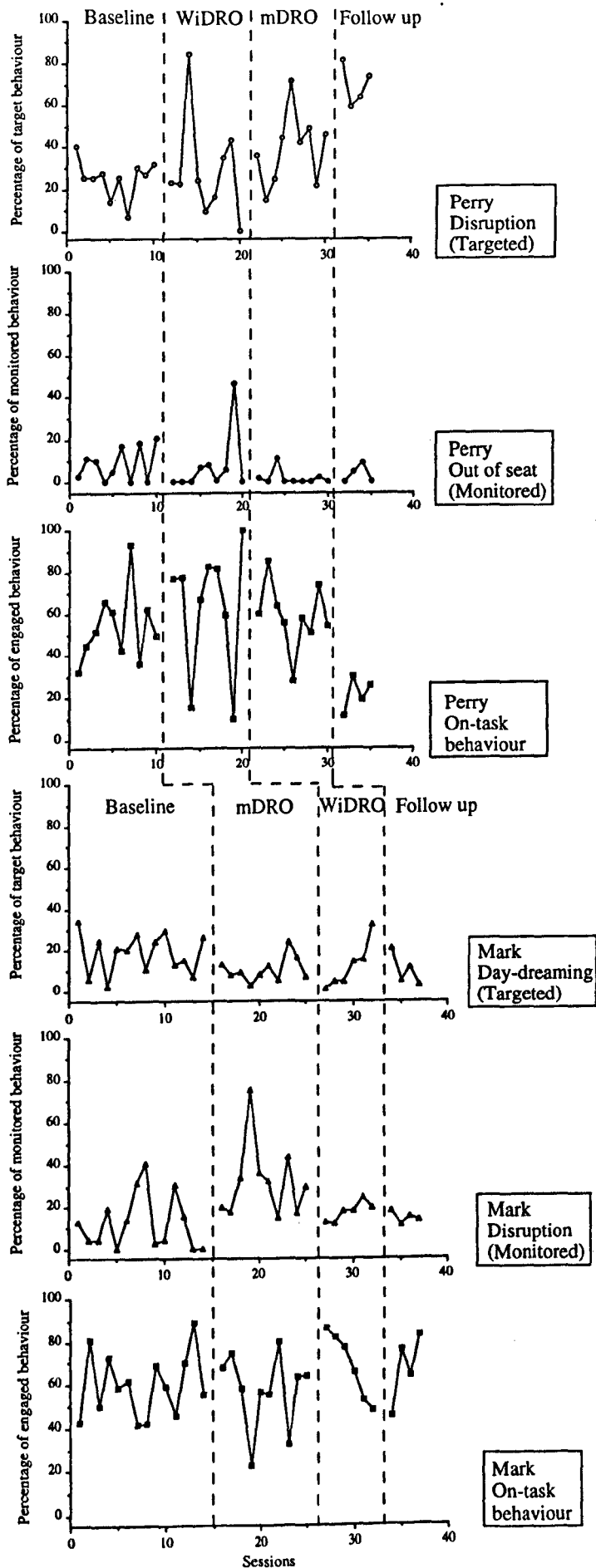
Figure 4.1 : The percentages of the duration of the target and monitored behaviours.
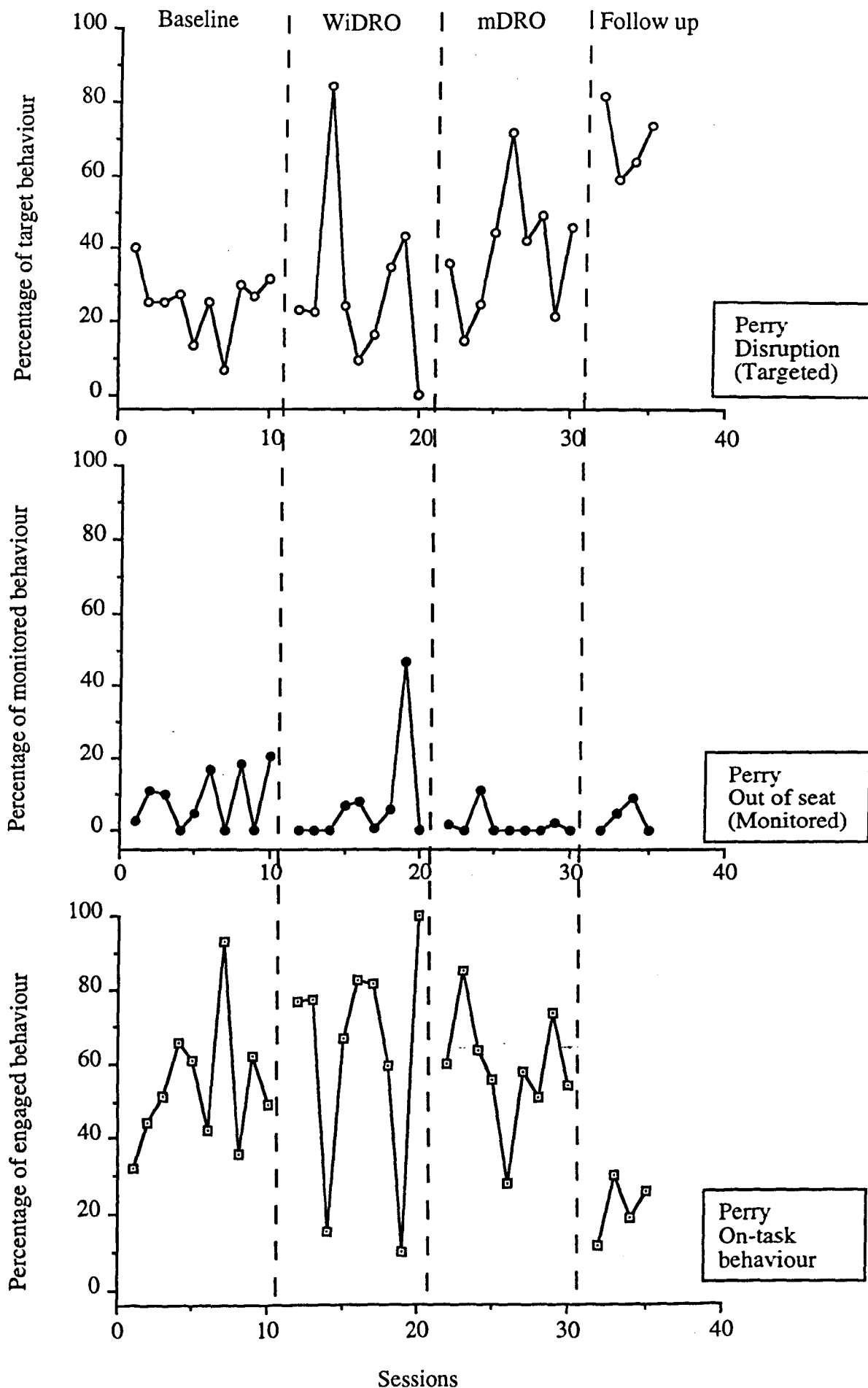
Figure 4.2 : The percentage of the duration of Perry's behaviours

73
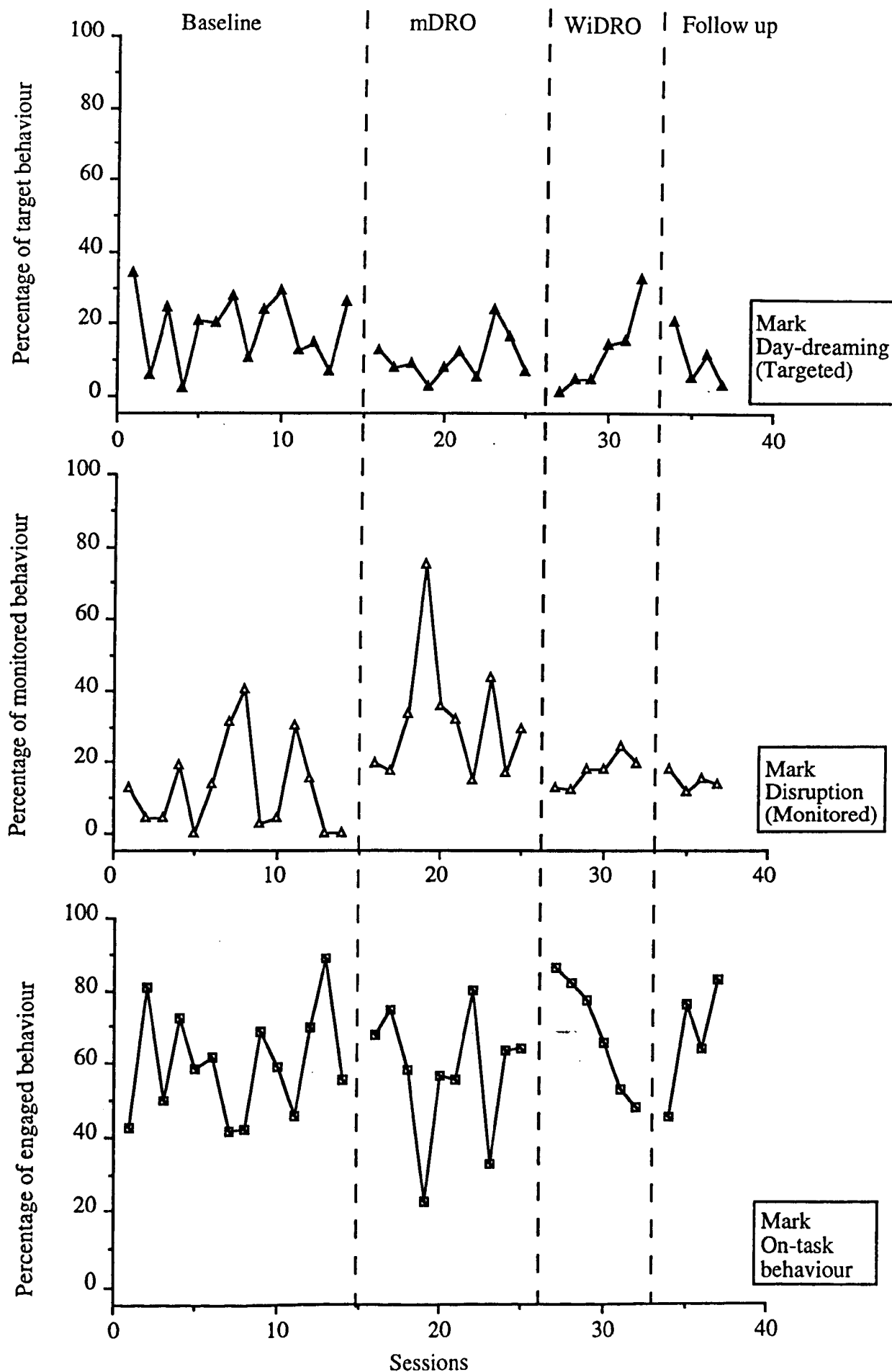
Figure 4.3 : The percentage of the duration of Mark's behaviours.

74

Mark's target behaviour, 'day-dreaming', averaged 18% in baseline, and ranged from 2% to 34%. During momentary DRO, 'day-dreaming' averaged at 10% and ranged from 3% to 24%. During whole-interval DRO, 'day-dreaming' averaged at 12% and ranged from 1% to 32%.

Mark's monitored behaviour, disruption, averaged at 13% in baseline and ranged from 0% to 40%. During momentary DRO, disruption averaged at 32% and ranged from 15% to 75%. During whole-interval DRO, disruption averaged at 17% and ranged from 12% to 24%. Mark's on-task behaviour averaged at 60% in baseline and ranged from 41% to 89%. During momentary DRO, on-task behaviour averaged at 57% and ranged from 22% to 80%. During whole-interval DRO, on-task behaviour averaged at 68% and ranged from 48% to 86%.

During the follow-up sessions Mark's 'day-dreaming' averaged at 10% and ranged from 3% to 21%. Disruption averaged at 15% and ranged from 12% to 18%. On-task behaviour averaged at 67% and ranged from 45% to 83%.

DISCUSSION

Both schedules increased one subject's behaviour from baseline, with momentary DRO producing a larger increase than whole-interval DRO. For the other subject, both schedules reduced the target behaviour from baseline, with momentary DRO producing a slightly larger reduction than whole-interval DRO. These results will now be considered in more detail.

Both schedules increased Perry's targeted behaviour (disruption) from baseline. In whole-interval DRO there were extreme fluctuations of disruption, with a low of zero and a high of 85%. That single high point at session 13 was the main reason why the mean rate of behaviour during whole-interval DRO was higher than in baseline. Other than that one

session the disruption in whole-interval DRO was similar to that during baseline. Disruption during momentary DRO did not reach such high peaks, but it contained a larger number of sessions in which disruption lasted for long periods. In no session during momentary DRO was disruption lower than 15%. The high point of disruption in whole-interval DRO coincided with the first reliability check. This meant that there were two people observing Perry instead of just one, and this seemed to inspire him to play up and disrupt his classmates more than previously. However, Perry's disruption seemed to increase most noticeably shortly after Mark's intervention began. As Mark began to receive reinforcers, Perry appeared to become annoyed and his behaviour worsened accordingly. A competition started up between the two boys, consisting of them boasting about the number of stars they had received and repeatedly adding up the totals on their star-charts. The stars seemed to become secondary to the competition between the boys, and ultimately it is possible that the stars became aversive to Perry as they had become associated with taunts and boasting between Mark and himself. As with Experiment 1, it is worth noting that it was possible for Perry to earn twice as many reinforcers as Mark, as Perry's DRO interval length was half the length of Mark's. It is not clear if this factor contributed to the competition between the boys. However, the problem of competitiveness highlights a real difficulty in administering treatment programmes to children in the same class, and may have implications for employing similar interval lengths to avoid discrepancies in reinforcer level.

There were also external events which may have contributed to the fluctuations in Perry's behaviour. At session 18, his teacher informed him that his mother would soon be coming to the school to hear him read, and told him that if he didn't improve his work his mother would not listen to him read. In the session after this, Perry's disruption fell to zero, apparently as a result of the teacher's threat. After his mother had been to the school and listened to him read, Perry's disruption again rose to high levels (sessions 23 and 24), possibly as an aftermath of the excitement of her visit. During this same period, Perry became aggressive towards others, including the experimenter and Mark. Thus it appears

that environmental events exerted far greater control over Perry's behaviour than the DRO schedules.

There was some covariation between Perry's targeted behaviour and the monitored behaviour of 'out of seat'. Figure 4.2 shows that where disruption was high, 'out of seat' behaviour tended to be low, and vice versa. However, apart from a high point at session 18, 'out of seat' behaviour during the intervention tended to be lower than baseline. It is not clear if this is because the schedules had a suppressive effect on the monitored behaviour, or if it was simply because disruption was generally higher during intervention and so the covarying 'out of seat' behaviour did not occur so much. It may also be that Perry misunderstood the requirement that he should not disturb anyone, and felt that as long as he was sitting in his seat he wasn't disrupting his classmates. Certainly it was not clear that he understood the connection between not disrupting and receiving stars, even though this connection was explained to him.

There was, unsurprisingly, a high level of covariation between Perry's disruption and his on-task behaviour; presumably the two behaviours were incompatible. Figure 4.2 shows very clearly that when disruption was low, on-task behaviour was high; disruption therefore interfered with Perry's own learning as well as that of others.

During the follow-up sessions, three months later, Perry's disruption rose to the highest levels during the experiment. One explanation for this is that Perry had come to find the experimental situation and, by association, the experimenter, aversive. Even though there was no intervention during the follow-up, the simple presence of the experimenter may have been enough to cause Perry's behaviour to deteriorate. It is also possible that between the end of the intervention and the follow-up Perry's behaviour had become more disruptive for external reasons, and the follow-up merely recorded ongoing high levels of disruption. Perry's results do not allow a statement about treatment generalisability to be made, as there was no suppression of behaviour during intervention.

Mark was not adversely effected by the schedules as Perry was. The means show that both schedules reduced his targeted behaviour, although neither showed very large reductions. Momentary DRO reduced the variance in levels of disruption, and apart from one or two slightly higher points towards the end of the intervention, showed a fairly consistent low rate of behaviour. Whole-interval DRO began by suppressing the behaviour to levels below that of momentary DRO, but by the fourth session showed an increase which continued until the end of the intervention. The last session in whole-interval DRO showed levels of day-dreaming almost as high as the highest levels during baseline. It is likely that this upward trend was due to the teacher informing Mark that the experimenter would be leaving the following day. This information was given at session 27, and it is from that point that day-dreaming started to increase. The teacher confirmed that Mark appeared to feel a rapport with the experimenter, and she considered that Mark's behaviour deteriorated when he heard that the experimenter would be leaving. Thus once again the pattern of results may have been more influenced by external events than by the DRO schedules. During follow-up, Mark's day-dreaming was at levels similar to or just below baseline levels. Mark's follow-up results therefore do not provide information regarding the generalisability of treatments, as much of the variability in his behaviour must be attributed to external events and not to the interventions.

There did not seem to be a consistent relationship between Mark's targeted behaviour and his monitored behaviour. Sometimes when one was low the other was high (e.g. session 18), but on other occasions, both behaviours rose in tandem (such as during whole-interval DRO). Mark's disruption was particularly high during momentary DRO. This may be in part because at the beginning of this intervention, Mark realised that in order to get a star he simply had to avoid day-dreaming. He told the experimenter that as long as he didn't stare into space he would get a star, so this meant he could speak to others in the classroom. At session 19 the teacher told Mark that as well as not day-dreaming he should not interrupt others but should get on with his own work. This did have a temporary reductive effect on

disruption but it still remained fairly high for the duration of momentary DRO. Another possibility is that during the first intervention, Mark was more susceptible to Perry's taunts about the number of stars he was receiving; as time went on and Mark began to earn more stars he appeared to become less effected by this. Thus it may be that during momentary DRO Mark was more disturbed by Perry than later on, and his disruptive behaviour increased in response to this. However, there was a distinct contrast between disruption in momentary DRO and whole-interval DRO, which suggests that the increased levels of disruption may have had something to do with the schedule itself. For example, the frequent reinforcement in momentary DRO might have led Mark to show his friends the star chart more frequently, or might have made him excited so that he was more inclined to get up and interact with others.

There was clearer response covariation between disruption and on-task behaviour, and to a lesser extent between day-dreaming and on-task behaviour. Overall levels of on-task behaviour did not vary greatly across the intervention, but different patterns could be clearly seen. In baseline, where disruption was high, on-task behaviour was low, and vice versa (this pattern is particularly clear at sessions 7 and 8, and again at 13). A similar, though less striking covariation occurred between day-dreaming and on-task behaviour. During momentary DRO there was a very clear relationship between disruption and on-task behaviour: the two patterns were almost mirror images of each other. At session 18, for example, where disruption reached its highest point, on-task behaviour fell to its lowest level. A similar relationship was not seen between day-dreaming and on-task behaviour in momentary DRO. However, during whole-interval DRO day-dreaming rose and on-task behaviour fell in almost identical opposing patterns. Day-dreaming in whole-interval DRO began at almost its lowest point during the study and then rose to a level as high as the highest points during baseline. Meanwhile, on-task behaviour began at almost its highest point and then fell to a level similar to the lowest points during baseline. The link between the two behaviours in this intervention is unmistakable. A similar, though far less pronounced pattern was seen between disruption and on-task behaviour. During follow-

up, the relationship again appeared to be between day-dreaming and on-task behaviour; again, the patterns were mirror-images of each other.

Mark's results, then, displayed a complex relationship between the three behaviours which were observed in this study. Sometimes there was a negative relationship between on-task behaviour and disruption, and sometimes between on-task behaviour and day-dreaming. Sometimes there was a negative relationship between day-dreaming and disruption, and sometimes a positive relationship. This demonstrates the shifting nature of behaviour and the difficulty of making conclusive predictions about the effect of altering one behaviour on another.

Occasionally, it seemed that the boys' behaviour was influenced by each other. For example, at session 18, Mark's disruption reached its highest point, and Perry's disruption was also at a high point here. It was in this session that Perry was told that his mother would not hear him read if he didn't behave. It is possible that the tension surrounding this event communicated itself to Mark. On other occasions, however, there was not a clear correspondence between the boys' behaviour.

The confounding effect of external environmental events appears to have played a major part in the behaviour of both subjects. The sessions took place in a classroom with many other children present, and there was always a great deal of activity going on. Because of these extraneous events, it is difficult to draw conclusions about the DRO schedules in the present experiment. What is clear is that one subject appeared to find the process somewhat aversive; that subjects became competitive when they saw the other receiving reinforcers; and that there was a great deal of response covariation between behaviours, often in very complex forms. The fact that the subject for whom the schedules were effective received full instructions and the other subject did not suggests the value of providing such instructions. However, there were attendant problems in providing full

instructions to Perry, and it is by no means certain that this omission played any part at all in the poor performance of the schedules in his case.

What is clear is that giving Mark full instructions about momentary DRO, including giving him a signal indicating the end of the interval, did not noticeably interfere with the effectiveness of this schedule. A key difference between the instructions given to Mark and those given by Repp et al. (1983) is that the present study did not describe what would happen if target behaviour *was* displayed at the signal. Repp et al. told their subjects, "If you are being disruptive [when Ms Smith raises her hand] you will not earn a treat". In the present study subjects were only told that they would earn a reinforcer if they were refraining from the target behaviour at a given moment or signal. It may be that the additional information provided by Repp et al. helped to clarify the significance of the single moment and thus weakened the control of momentary DRO. Further experimentation would be necessary to confirm this.

It is already known that whole-interval DRO can be effective with non-learning disabled subjects (Barton & Barton, 1978; Christensen & Sanders, 1987). This experiment showed that the same is also true of momentary DRO, as the subject for whom both schedules worked did not have learning disabilities.

To sum up: the mixed findings of this experiment showed that for one subject, both schedules were effective in reducing disruptive behaviour, with a very slightly greater reduction in momentary DRO. For the other subject, both schedules were ineffective in reducing his behaviour; and this shows that like whole-interval DRO, momentary DRO can sometimes increase target behaviour. However, these conclusions are somewhat weakened by the unknown effects of extraneous variables which appeared to distort the pattern of behaviour seen in the two subjects. These variables mean that the follow-up measures cannot provide any information regarding the generalisability of treatments over time.

GENERAL DISCUSSION
EXPERIMENTS 1, 2, 3 & 4

The four experiments outlined in this chapter provide strong support for the hypothesis that momentary DRO can be as effective as whole-interval DRO in the reduction of inappropriate behaviour. Overall, there were seven subjects. Momentary DRO was considerably more effective than whole-interval DRO for one subject; and as effective for five. For the seventh subject, both schedules increased the behaviour from baseline, though momentary DRO increased it to a greater extent. These findings support those of Derwas and Jones (1993) that momentary DRO can be as or more effective as whole-interval DRO. They fail to support the findings of Repp et al. (1983) and Barton, Brulle and Repp (1986) that momentary DRO is weaker than whole-interval DRO and is of use merely to continue suppression begun by whole-interval DRO.

These experiments tested several of the aims outlined in 'Aims of the present thesis' at the end of Chapter 1. They examined two different types and rates of behaviour: high-rate stereotypy and lower-rate disruptive behaviour. Experiments 1 and 2 worked with high-rate stereotypy. Experiment 1 found that momentary DRO was more effective than whole-interval DRO for one subject and as effective for the other. The subject for whom it was more effective received greater levels of reinforcement in momentary DRO, while the subject for whom both schedules worked equally received similar amounts of reinforcement in the two schedules. This suggests that momentary DRO's effectiveness is linked to the greater density of reinforcement it provides. This issue is explored further in the next chapter. Experiment 2 found both schedules to be effective, although momentary DRO produced slightly greater reductions. The DRO schedules also had a highly suppressive effect on two monitored inappropriate behaviours. It was suggested that the results in this experiment were due more to the effects of experimenter attention than the tangible reinforcer. The attention feature of DRO schedules is considered further in the next chapter. Experiments 1 and 2 fulfilled another aim of this thesis, in that they replicated

Derwas and Jones's findings that momentary DRO can be as or more effective as whole-interval DRO in reducing high-rate stereotyped behaviour.

Experiments 3 and 4 worked with lower-rate disruptive behaviour. Experiment 3 showed that, contrary to expectations, momentary DRO could be as effective as whole-interval DRO in reducing such behaviour. Experiment 4 produced uneven results, but essentially showed that where momentary DRO failed to work, whole-interval DRO failed to work also. Thus these findings show that momentary DRO can have more control over lower-rate behaviours than has previously been reported. Further experiments will be necessary to determine if momentary DRO can be effective with much lower rate behaviours. It may be, for example, that momentary DRO cannot control behaviours which occur less than once an hour, or less than three times a day.

Experiments 3 and 4 failed to support the hypothesis of Repp et al. that momentary DRO would not be effective for people who could readily discriminate between the two schedules. Three subjects had only mild learning disabilities and the fourth subject did not have learning disabilities. The distinction between schedules was explained and subjects did not appear to exploit the opportunity for disruption during intervals allowed by momentary DRO. This may be because they did not understand the instructions. They may have assumed that they still had to behave appropriately for the whole interval in order to receive reinforcement. Or it may have been that because less emphasis was placed on the connection between a single moment and a reinforcer, or on a signal and a loss of a reinforcer, subjects felt they should not risk the possibility of disrupting in case they lost a reinforcer. In whole-interval DRO, however, if subjects disrupted at all they knew they had lost the reinforcer for that interval and so there was little point refraining from disruption immediately afterwards. Alternatively, subjects may have fully understood the implications of the instructions, but the greater levels of reinforcement and/or attention in momentary DRO might have meant that they had less need of the alternative reinforcement obtainable from disruption.

Momentary DRO, then, can be used with capable subjects, but the necessary parameters are not clear. A level of ambiguity in instructions may be necessary to ensure that the inappropriate behaviour does not simply dip at a signal and then rise again. The relative effectiveness of momentary DRO and whole-interval DRO with verbal subjects is considered further in Chapter 4.

Experiment 4 aimed to look at the generalisability over time of the treatments, via a follow-up taken three months later. However, due to extraneous variables and ambiguous results during intervention, a statement about generalisability cannot be made. Further experiments would be necessary to look at this aspect of momentary DRO.

Experiments 3 and 4 showed that momentary DRO can be practical to use in the classroom, although measures should be taken to avoid competitiveness between classmates. For example, different types of reinforcers could be used for each subject, or it should be ensured that each subject has similar opportunities for reinforcement. Informal evidence from these two experiments showed that a non-resetting whole-interval DRO schedule was easier to implement than a resetting interval when recording duration of behaviour.

Overall, these experiments showed that momentary DRO can be as effective as whole-interval DRO under a variety of different conditions with a variety of different subjects. In some cases it can be more effective. Two of the questions which were raised in this chapter concern the role played in momentary DRO by greater levels of reinforcement, and the effect of increased attention. These factors are considered in the next chapter.

Chapter 3

The effect of increased reinforcement in momentary DRO

It is clear from the previous experiments that in some circumstances momentary DRO schedules can be as - or more - effective than whole-interval DRO schedules. One possible reason for this is that in momentary DRO, the client receives more reinforcement than in whole-interval DRO. In a whole-interval DRO schedule, the timing of the interval stops if the inappropriate behaviour occurs, and does not start again until the behaviour has ceased. However, in a momentary DRO schedule, the time interval is worked through without stopping and the next interval starts directly the previous interval is complete. This means that there are more opportunities for reinforcement in momentary DRO. In addition, in whole-interval DRO every occurrence of the target behaviour makes reinforcement less likely, whilst in momentary DRO the target behaviour does not limit the possibility of reinforcement unless it occurs at one particular moment. Both these components of momentary DRO allow a greater density of reinforcement. Experiment 1 confirmed that overall, a momentary DRO schedule produced significantly higher levels of reinforcement than whole-interval DRO. This factor was associated with differences in schedule effectiveness.

The experiments in this chapter were designed to explore the issue of increased reinforcement in momentary DRO.

EXPERIMENT 5

A comparison of the effectiveness of momentary DRO, whole-interval DRO, and momentary DRO using half-length intervals

INTRODUCTION

In order to investigate the effect of greater density of reinforcement, an experiment was carried out to compare the effect on behaviour of standard levels of reinforcement in momentary DRO with a momentary DRO schedule that provided far greater levels of reinforcement. This experiment worked with a single subject, and employed an alternating treatments design. The subject displayed extremely high-rate stereotypy, and because of the intensity of concentration that was required during schedule implementation, a record was not kept of the amount of reinforcement provided.

The aims of this experiment were as follows:

1.   To compare the effectiveness of whole-interval DRO and momentary DRO with a momentary DRO schedule that provides higher levels of reinforcement.

2.   To explore the relationship between schedule effectiveness and density of reinforcement.

METHOD

Subject

Derek was 40 years old and had severe learning disabilities. He was a resident at a hospital for people with learning disabilities. He was taking a variety of medication, including Haloperidol and Orphenadrine. He had no self-help skills and no language, although he was able to understand simple instructions.

Response definition

Derek had a number of inappropriate behaviours, including rocking, hand-mouthing, and occasional aggression towards others. It was decided after consultation with care-staff that Derek's highest-rate stereotypy - rocking - would be targeted in this experiment. This behaviour was operationally defined as moving his whole torso vigorously back and forth whilst in a sitting position. It was also decided to monitor Derek's second highest rate stereotypy, hand-mouthing. This behaviour was operationally defined as rubbing his hands or fingers against his lips or teeth, or inside his mouth.

Apparatus

A personal cassette-player and headphones, pre-recorded audio-tape, recording sheets and stop-watch were used in this experiment, as described in the Apparatus section for Experiment 1 (Chapter 2).

Design

An alternating treatments design was employed in this experiment. The baseline phase was followed by an intervention phase in which three treatments were randomly alternated. Each five minute session was immediately followed by the next, according to a pre-determined list of treatments. The treatments were:

(i)   Whole-interval DRO

(ii)   Momentary DRO

(iii) Momentary DRO using half the length of the interval used in the first two schedules (Momentary DRO/2).

This last schedule has not been used before. It was designed for this experiment as a schedule which would operate on the same principles as momentary DRO but allow a far greater opportunity for reinforcement. To measure response covariation effects, Derek's second highest-rate stereotypy was monitored but not consequated. The experiment took place over two months.

Reinforcer assessment

Four care-staff completed Motivation Assessment Scales. All agreed that Derek's rocking was maintained by sensory stimulation. Tangible reinforcement was also rated very highly. As care-staff could not suggest any sensory stimuli which Derek might like, a variety of edible items were selected, and a multiple-choice reinforcer assessment was carried out over six trials. Following this procedure, chocolate was found to be chosen more than any other item, and this was therefore used during the experiment.

Inter-response time

The length of time between occurrences of Derek's rocking was recorded over several days. The average length of time between occurrences was 5.96 seconds. Therefore, the DRO interval length was set at 5 seconds.

Procedure

Baseline

Baseline measures of the target behaviour were taken until stability was achieved.

Intervention

During intervention, Derek was randomly presented with three different reinforcement schedules.

Whole-interval DRO

During this intervention Derek was given a reinforcer if he had not rocked at all during the interval of 5 seconds. Timing of the DRO interval was suspended while he ate the reinforcer. If he did rock during a interval, timing of that interval ceased, and the stop-watch was reset. A new interval began when the target behaviour ceased. Thirteen sessions of this intervention were carried out.

Momentary DRO

During this intervention Derek was given a reinforcer if he was not rocking at the moment between two 5-second intervals. Timing of the DRO interval was suspended while he ate the reinforcer. If he was rocking at that moment, no reinforcer was given, and a new interval started. Twelve sessions of this intervention were carried out.

Momentary DRO/2

During this intervention the DRO interval was set to half the interval used in the other two schedules, that is, to 2.5 seconds. Derek was given a reinforcer if he was not rocking at the moment between two 2.5-second intervals. The rest of the procedure was as for momentary DRO. Thirteen sessions of this intervention were carried out.

Inter-observer agreement

For 12% of the sessions, observations were taken by a second observer. Agreement between raters was calculated as for Experiment 1. Agreement for Derek's rocking ranged from 90% to 100% with a mean of 94%. Agreement for his hand-mouthing ranged from 83% to 100% with a mean of 93%.

RESULTS

Figure 5.1 shows the percentage of observations in which the targeted and monitored stereotyped behaviours occurred. In baseline, rocking ranged from 66% to 100%, averaging 86%. During whole-interval DRO, rocking ranged from 31% to 100%, averaging 79%. Rocking during momentary DRO ranged from 17% to 100%, with a mean of 66%. In momentary DRO/2, rocking ranged from 13% to 97%, with a mean of 68%.
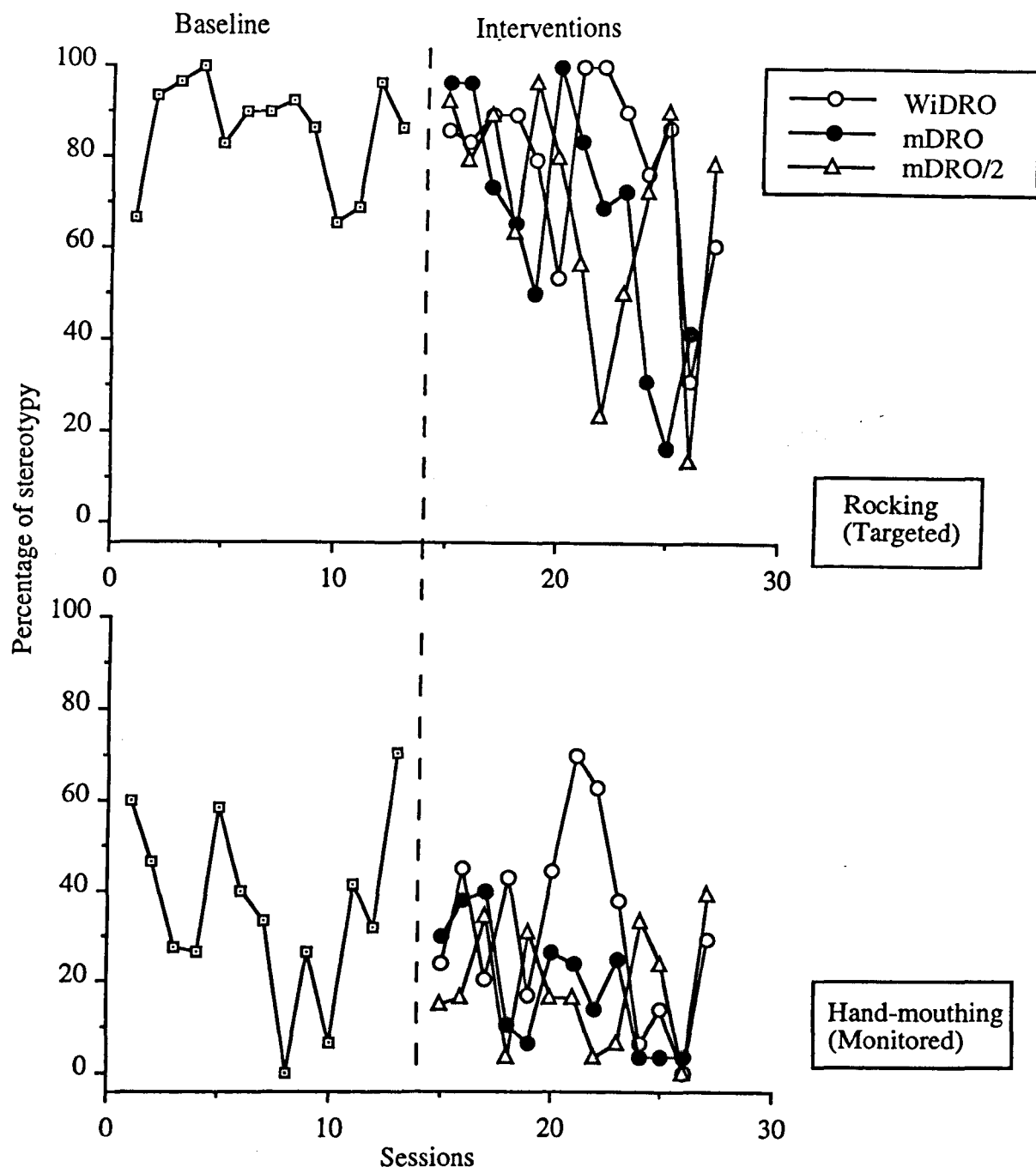
Figure 5.1 : The percentage of observations in which Derek's targeted and monitored behaviours occurred.

In order to assess whether there was a statistically significant difference between the schedules, an approximation to the randomization test was carried out. This took the form of a one-way analysis of variance. The rationale behind this test is given in Appendix B.

The ANOVA showed that there was not a significant difference between any of the three schedules (F(2,35) = 0.99, p>0.1).

The monitored stereotyped behaviour, hand-mouthing, was at its highest during baseline (mean = 36%, range = 0% to 70%). It was at its lowest during momentary DRO/2 (mean = 8.5%, range = 0% to 39%), and hardly any higher during momentary DRO (mean = 19%, range = 3% to 40%). It rose back to almost baseline levels during whole-interval DRO (mean = 32%, range = 0% to 70%). An analysis of variance showed that the rate of monitored behaviour differed across schedules, although this difference did not reach the 5% level of significance (F(2, 35) = 2.839, p<0.1). However, Fishers PLSD post-hoc tests showed that there was a significant difference between whole-interval DRO and momentary DRO/2: monitored behaviour was significantly higher under whole-interval DRO than under momentary DRO/2. Whole-interval DRO did not differ significantly from momentary DRO, and the two momentary DRO schedules did not differ significantly from each other.

DISCUSSION

The results of this study show that although none of the treatments reduced the target behaviour substantially, momentary DRO and momentary DRO/2 were more successful than whole-interval DRO. However, the ANOVA showed that there was not a statistically significant difference between schedules. Figure 5.1 shows that the pattern of rocking under both whole-interval DRO and momentary DRO was a downwards trend, with momentary DRO at a lower level than whole-interval DRO. It may be that if the experiment had been continued, rocking would have been more substantially reduced. Rocking under momentary DRO/2 had shown a downward trend earlier but had then risen again.

As it was possible for Derek to receive the reinforcer at a very high frequency, it was likely that he quickly became sated. This may have been the reason why none of the schedules

were highly effective. Satiation may also have caused the lack of difference between momentary DRO and momentary DRO/2, as the extra reinforcement in momentary DRO/2 would not have had any impact. Therefore, any advantage of the increased reinforcement in momentary DRO/2 would not have been seen for this subject. This is obviously one of the problems of using edible reinforcers, although other, less tangible reinforcers have also been found to be prone to satiation (Vollmer & Iwata, 1991). Future studies working with such high-rate behaviours should make use of a variety of different reinforcers so that satiation does not occur so quickly. They should also incorporate an escalating schedule so that as soon as the subject achieves criterion at the short DRO interval, a longer interval can be implemented. This would make satiation less likely.

Although there may have been a problem of reinforcer satiation, the difference between whole-interval DRO and the two momentary DRO schedules suggests that the greater opportunities for reinforcement in momentary DRO had a positive effect on this subject. Because the behaviour occurred so frequently, it would have been difficult for Derek to earn high levels of reinforcement during whole-interval DRO, as any instance of the behaviour would reset the interval. Although a record was not kept of the relative levels of reinforcement, the nature of the schedules would strongly suggest that there were far higher levels in momentary DRO and momentary DRO/2 than in whole-interval DRO. The reasons why higher levels of reinforcement might produce greater reductions in target behaviour are not straight-forward and may be concerned less with the actual stimulus used than the changes in environment which programmed reinforcement brings about. This is considered in more detail in the general discussion, below.

Both momentary DRO interventions reduced Derek's monitored behaviour, hand-mouthing. The post-hoc tests showed that there was a significant difference in levels of this behaviour under momentary DRO/2 and whole-interval DRO. There was almost as large a difference between momentary DRO and whole-interval DRO. It is likely that this behaviour diminished when the levels of reinforcement were high, because the reinforcers

were chocolate. This meant that Derek's mouth was occupied while levels of reinforcement were high and thus the alternative oral stimulation of hand-mouthing occurred at a lower rate. Levels of hand-mouthing during whole-interval DRO were little different from those during baseline, suggesting again that lower levels of programmed reinforcement were provided during this schedule. An alternative explanation, however, is that rocking and hand-mouthing tended to occur together, and as rocking was reduced, hand-mouthing was also reduced. The graphs provide some support for this explanation. Although the two behaviours appeared to co-vary during baseline, under all three interventions both behaviours followed similar patterns to each other. When one was high the other tended to be high, and when one was low the other tended to be low. Informal observation also supports this, as the two behaviours did tend to occur simultaneously.

It was felt that there had not been a clear enough distinction between the three interventions, as they were rapidly alternated without a break between them. It is possible that the effects of one intervention 'leaked' into the next; and there might have been a clearer distinction seen between momentary DRO and momentary DRO/2, had sessions been delineated in some way. It was also felt that it was necessary to keep a record of the amount of reinforcement provided in the three schedules. It was therefore decided to replicate this experiment in order to examine the effect of momentary DRO/2 with these two factors taken into account.

EXPERIMENT 6

A comparison of the effectiveness of momentary DRO, whole-interval DRO, and momentary DRO using half-length intervals: A further development of Experiment 5

INTRODUCTION

There were two main differences in methodology between Experiment 5 and the present experiment. The first was that during the alternation of treatments, a five-minute break was taken after each session. This was to make a clearer distinction between the three treatments. The second difference was that a record was kept of the number of reinforcers delivered in each of the three schedules, to allow a comparison to be made.

METHOD

Subject

Keith was 33 years old and had severe learning disabilities. He was a resident at the same hospital as Derek. He was taking a variety of medication, including Haloperidol, Procylidine and Carbomazepine. He had some limited self-help skills and had no language, although he seemed able to understand simple instructions.

Response definition

Keith displayed a variety of inappropriate behaviours, including head-slapping, hand-biting, hand-clapping, skin-picking and spinning on his heels. After initial observation sessions and consultation with care-staff, it was decided that Keith's highest-rate stereotypy - skin-picking - would be targeted in this experiment. This behaviour was operationally defined as touching, rubbing or scratching at the base of his nose (between

the nostrils) or between the base of the nose and the top lip. It was also decided that his second highest-rate stereotypy - 'spinning' - would be monitored. This behaviour was operationally defined as standing and turning round 360° more than once, either on the spot or while moving round the room.

Apparatus

Apparatus used in this experiment was the same as for Experiment 5.

Design

The design of this experiment was the same as for Experiment 5. To measure response covariation effects, Keith's second highest-rate stereotypy, spinning, was monitored but not consequated. The experiment took place over three months.

Reinforcer assessment

Four care-staff completed Motivation Assessment Scales. There was a high level of agreement between raters; three of the four rated Keith's skin-picking as most likely maintained because of the sensory stimulation it provided. A variety of sensory-type stimuli were tested, but Keith seemed averse to them and would move away. Finally, a variety of edible items were selected after consultation with care-staff, and a multiple-choice reinforcer assessment was carried out over six trials. Following this procedure, Bakewell tart (a type of cake) was found to be chosen more than any other item and this was therefore used during the experiment.

Inter-response time

The length of time between occurrences of Keith's skin-picking was recorded over several days. The average length of time between occurrences was 38 seconds. The DRO interval length was therefore set at slightly less than this (30 seconds).

Procedure

The procedure was the same as for Experiment 5, other than the changes outlined below.

Intervention

A small gap of five minutes between the presentation of each treatment was implemented, in an attempt to make a clear distinction between each treatment.

Whole-interval DRO

During this intervention Keith was given a reinforcer if he had not skin-picked at all during the interval of 30 seconds. Timing of the DRO interval was suspended while he ate the reinforcer. If he did skin-pick during an interval, timing of that interval ceased, and the stop-watch was reset. A new interval began when the target behaviour ceased. Ten sessions of this intervention were carried out.

Momentary DRO

During this intervention Keith was given a reinforcer if he was not skin-picking at the moment between two 30-second intervals. Timing of the DRO interval was suspended while he ate the reinforcer. If he was skin-picking at that moment, no reinforcer was given, and a new interval started. Ten sessions of this intervention were carried out.

Momentary DRO/2

During this intervention the DRO interval was set to 15 seconds. Keith was given a reinforcer if he was not skin-picking at the moment between two 15-second intervals. The rest of the procedure was as for momentary DRO. Ten sessions of this intervention were carried out.

Inter-observer agreement

For 16% of the sessions, observations were taken by a second observer. Agreement between raters was calculated as for Experiment 1. Agreement for Keith's skin-picking

ranged from 47% to 83% with a mean of 73%. Agreement for his spinning ranged from 83% to 100% with a mean of 97%.

RESULTS

Figure 6.1 shows the percentage of observations in which the targeted and monitored stereotyped behaviours occurred. In baseline, skin-picking ranged from 0% to 41%, averaging 19%. During whole-interval DRO, this behaviour increased, with a range of 0% to 50%, averaging 30%. Both momentary DRO and momentary DRO/2 were highly effective in reducing the stereotypy. Skin-picking during momentary DRO ranged from 0% to 20%, with a mean of 12%. In momentary DRO/2, the range was broader (0% to 30%) but the mean was lower (9%).

An approximation to the randomization test was carried out in the form of an ANOVA, as for Experiment 5. This showed that there was a highly significant difference between schedules ($F(2, 27) = 12.291$, $p<0.001$). Fishers PLSD post-hoc tests showed that target behaviour under whole-interval DRO was significantly higher than under both momentary DRO, and momentary DRO/2. There was not a significant difference in target behaviour between momentary DRO and momentary DRO/2.

The monitored behaviour, spinning, was at its highest during baseline (mean = 9%, range = 0% to 41%). It was at its lowest during momentary DRO (mean = 3%, range = 0% to 7%), and not much higher during whole-interval DRO (mean = 3.9%, range = 0% to 13%). However, it rose back to almost baseline levels during momentary DRO/2, although with a much lower range (mean = 8.8%, range = 0% to 20%). An analysis of variance showed that the rate of monitored behaviour differed across schedules ($F(2,27) =3.774$, $p<0.05$).
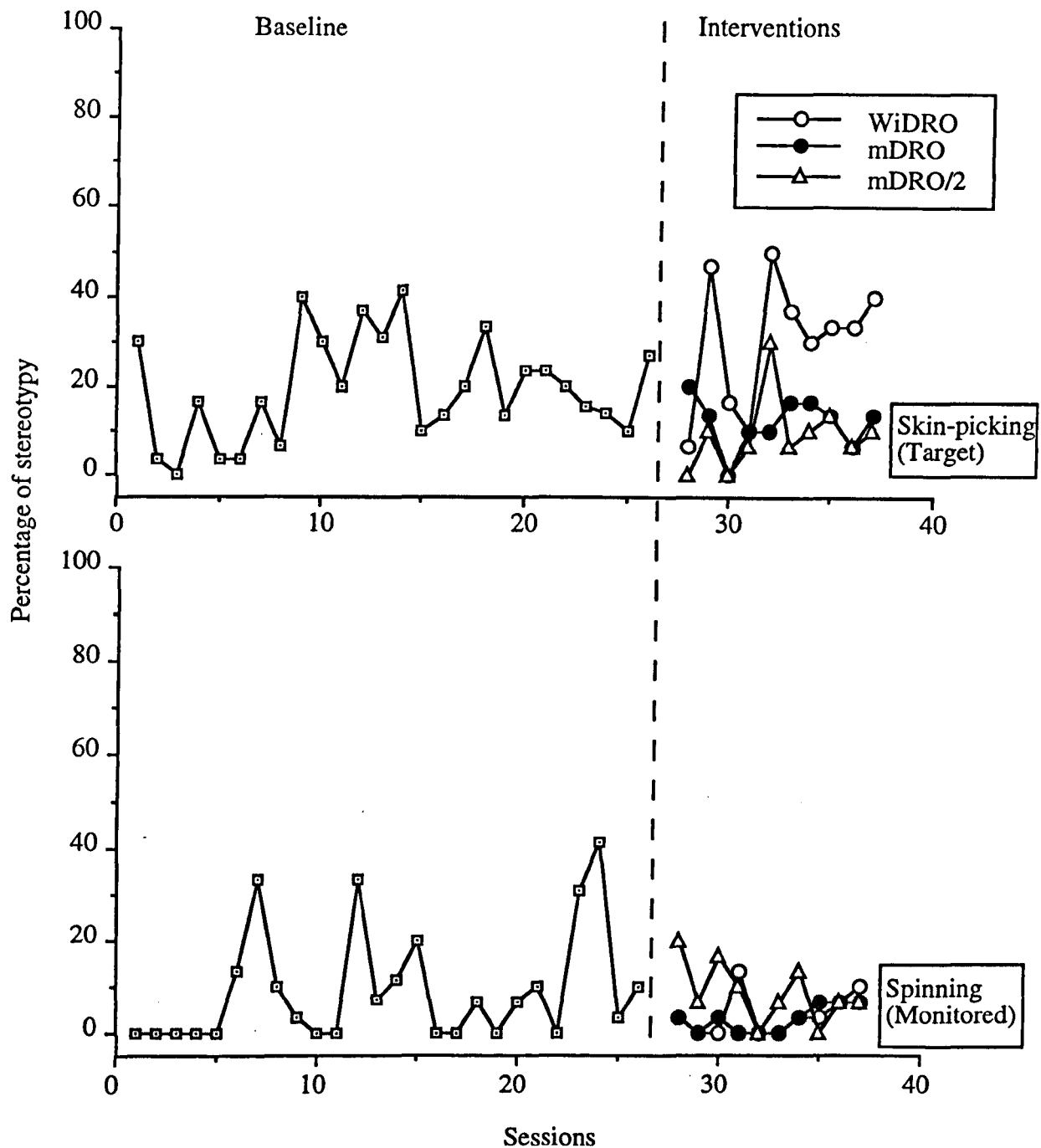
Figure 6.1 : The percentage of observations in which Keith's targeted and monitored behaviours occurred.

Fishers PLSD tests showed that monitored behaviour under momentary DRO/2 was significantly higher than under both momentary DRO and whole-interval DRO. There was

not a significant difference in monitored behaviour between whole-interval DRO and momentary DRO.

Number of reinforcers

The number of reinforcers delivered in each session were recorded, and the means, standard deviations and ranges in each DRO schedule are presented in the table below.

Table 6.1 Means, standard deviations and ranges of the number of reinforcers given in the three DRO schedules

	Whole-interval DRO	Momentary DRO	Momentary DRO/2
Mean number of reinforcers	2.8	8.8	14.2
Standard deviation	1.687	1.814	3.425
Range	0 - 5	4 - 10	10 - 19

Almost three times as many reinforcers were given in momentary DRO than in whole-interval DRO; and more than one and a half times as many in momentary DRO/2 than in momentary DRO. A one-way analysis of variance was carried out to determine if the amount of reinforcement differed significantly across schedules. The ANOVA was highly significant ($F_{(2, 27)} = 54.6$, $p < 0.001$), and post-hoc Fishers PLSD tests showed that each of the comparisons was statistically significant. In other words, whole-interval DRO differed significantly from momentary DRO; whole-interval DRO also differed significantly from momentary DRO/2; and momentary DRO differed significantly from momentary DRO/2.

DISCUSSION

The results of this experiment showed that momentary DRO and momentary DRO/2 were significantly more effective than whole-interval DRO in reducing Keith's targeted

stereotyped behaviour. Whole-interval DRO actually increased Keith's stereotypy. This experiment provides further evidence that in some cases momentary DRO is more effective than whole-interval DRO. The ANOVA did not show a significant difference between the momentary DRO schedules. However, Figure 6.1 showed that stereotypy under momentary DRO/2 was as low or lower than stereotypy under momentary DRO in all but one session.

However, the overall ranges and means of momentary DRO and momentary DRO/2 were fairly similar. It is possible that some kind of floor effect occurred; that is, that the behaviour was reduced as low as it could be, and increased reinforcement could not have reduced it further. If this was the case, then the extent of the advantage of momentary DRO/2 over momentary DRO could not be known. Similarly, it is possible that Keith reached a satiation point with regard to the reinforcers, such that any more could not have any effect.

Under whole-interval DRO Keith's skin-picking increased considerably. It is possible that the lower levels of reinforcement in whole-interval DRO had a negative effect on this particular subject. Keith earned zero reinforcers in one session, and the most he earned was five, whilst in momentary DRO he never earned less than four reinforcers in a session, and mostly it was nearer eight or nine. In momentary DRO/2, the lowest number of reinforcers earned in a session was ten. One might expect that the low reinforcement in whole-interval DRO would have had no effect on the target behaviour, but as it increased it beyond baseline this suggests that there was even less reinforcement during whole-interval DRO than there was in the natural environment during baseline. There is some evidence for this. During baseline, when the experimenter was simply observing Keith, care-staff interacted normally with him. However, during whole-interval DRO, the experimenter was interacting with Keith and care-staff did not interact with him in case they got in the experimenter's way. It is possible that the withdrawal of staff communication during

whole-interval DRO deprived Keith of a source of reinforcement, which was not replaced by the low levels of programmed reinforcers awarded during the intervention.

The effect of the interventions on the monitored behaviour suggests that there was a functional equivalence between this behaviour and the target behaviour. Although lower overall during all three interventions than baseline, spinning was significantly higher than the other two schedules during momentary DRO/2, when skin-picking was at its lowest. When skin-picking peaked during momentary DRO/2 (session 31), spinning fell to zero. A similar pattern occurred throughout whole-interval DRO; although initially both behaviours were low, as skin-picking rose under this schedule, spinning remained low. And in momentary DRO, as skin-picking rose and fell, spinning tended towards an opposite pattern. Care-staff completed the MAS for spinning as well as skin-picking: overall, sensory stimulation was rated as the most likely maintaining variable for Keith's spinning as well as for his skin-picking. These results suggest therefore that both stereotypies fulfilled a similar function, and one occurred at higher rates to replace the other when it occurred at low rates.

GENERAL DISCUSSION
EXPERIMENTS 5 & 6

The momentary DRO/2 schedule was devised to allow investigation of a schedule which operated under the same parameters as momentary DRO, but which, due to its halved interval length, provided much higher levels reinforcement.

These experiments provide support for the suggestion that the increased reinforcement in momentary DRO can make it more effective than whole-interval DRO. In both experiments the momentary DRO schedules were more effective than whole-interval DRO, and Experiment 6 showed that they both provided far higher levels of reinforcement. In addition, momentary DRO/2 was slightly more effective than momentary DRO in Experiment 6, again suggesting the relevance of increased reinforcement. It is possible in both experiments that satiation occurred; and in Experiment 6 it is also possible that there was a floor effect with regard to how reduced the stereotypy could be, which meant that momentary DRO/2 simply could not achieve any more reductions because of external constraints.

There are several reasons why the greater reinforcement in momentary DRO may play an important part in its effectiveness. One reason is that if the reinforcer is functionally equivalent to that maintaining the inappropriate behaviour, then the greater reinforcement in momentary DRO might hasten the extinction process that takes place in such a case. As Mazaleski et al. (1993) have shown, extinction appears to be the most powerful component of DRO, and higher levels of reinforcement which are not contingent on the stereotypy will be likely to increase the rate of extinction.

In situations such as the present experiments, where the reinforcers used are arbitrary (e.g. not functionally relevant), extinction is unlikely to take place. This is because the relationship between the targeted behaviour and its reinforcing properties is not removed.

Although Mazaleski et al. (1993) doubted that a strong effect could occur in DRO with an arbitrary reinforcer, they did concede that a powerful arbitrary reinforcer could exert some control. It is possible, though, that even with an arbitrary reinforcer, some measure of extinction might occur. For example, a behaviour might be maintained by attention, but this might not be indicated by the functional analysis; if attention was provided along with the programmed reinforcer in a DRO schedule, then one might inadvertently extinguish the connection between the behaviour and attention. The role of attention is considered further below.

In cases where arbitrary reinforcers are used, and where there is no evidence that extinction has occurred, there are four possible reasons why the greater reinforcement in momentary DRO might cause this schedule to be more effective than whole-interval DRO.

The first reason is that the greater number of reinforcers means that an alternative reinforcer is competing at a high rate with the target behaviour (or on a 'rich' schedule, as O'Brien & Repp, 1990, term it). When the effect of the reinforcer starts to wear off, another reinforcer is likely to come along quickly in momentary DRO and distract the subject for a further time period. For subjects with severe learning disabilities and few resources, reinforcers presumably serve a dual function. They not only make a behaviour more (or in the case of DRO, less) likely; they also provide the subject with something to do other than the inappropriate behaviour, albeit for just a short while.

The second reason is related to this. It is possible that the greater reinforcers in momentary DRO might sate the subject so that s/he has less need for the alternative reinforcement of the inappropriate behaviour. Basically, greater external reinforcement may (temporarily) provide the subject with adequate levels of reinforcement, and thus contribute to reduced stereotypy.

The third reason is the speed of learning. If more reinforcers are provided then more clues are given as to what the subject is required to do in order to receive reinforcement. This may allow the subject to respond more speedily to the contingencies in operation.

The fourth reason is that increased reinforcement also means increased attention from the experimenter. It is possible that what the subject is actually responding to is attention (e.g. Brusca, Nieminen, Carter, & Repp, 1989; Jones & Baker, 1989). People with learning disabilities rarely get high levels of attention or interaction (Mazaleski et al. 1993) and often seek it in various ways. This is particularly likely in deprived or barren environments, such as the one where the experiments in this chapter were carried out. It is possible that in such situations the novelty of attention is more potent than any other stimulus. Although the functional analysis might be expected to show attention to be the most likely maintaining variable, care-staff may not associate the subject's inappropriate behaviour with a lack of attention as they themselves are so accustomed to the barren nature of the environment. This has implications for the validity of functional analyses, and not just checklists such as the Motivation Assessment Scale which are completed by care-staff. Analogue assessments, for example, present the individual with attention, toys, or other materials. In such cases the sheer novelty effect of such stimulation may be enough to produce high levels of alternative behaviour and concomitant low levels of inappropriate behaviour. It has been noted elsewhere (Toogood, 1996) that naturalistic observations of inappropriate behaviour record higher levels of behaviour than that recorded by analogue assessments. This suggests that the results of analogue assessments can be distorted by the effects of the unfamiliar environments in which clients are placed.

It should be possible to assess the effects of attention from baseline, because in baseline there is very little interaction between subject and experimenter. Therefore, any difference during intervention might be partly due to the interaction that comes when reinforcement is provided. However, there is an element of attention in baseline also: the experimenter sits close to the subject and watches them.

What can be concluded from the two experiments in this chapter is that the higher levels of reinforcement in momentary DRO contribute to its effectiveness. Particular support for this comes from the momentary DRO/2 schedule in Experiment 6, which produced greater reductions in behaviour than momentary DRO. Whether this is the sole reason for its effectiveness is considered in Chapter 7.

These experiments provided further evidence that for some individuals momentary DRO is more effective than whole-interval DRO. In addition, Experiment 6 showed that for some subjects whole-interval DRO increases the target behaviour from baseline. This confirms a possible negative side-effect of whole-interval DRO, noted by some other observers (Friman et al., 1986; Jones & Baker, 1988a): it can make an inappropriate behaviour increase.

Chapter 4

The effect of verbal rules on schedule responding

This chapter considers some of the issues concerning the effects of language and rules on the effectiveness of DRO schedules, an area which has been little considered (e.g. Homer & Peterson, 1980; Jones, 1991a). Some of the DRO research has been carried out with people who have language abilities (e.g. Barton & Barton, 1978), and some has been carried out with people who do not. Some studies have instructed the subjects as to the nature of the contingencies, and others have not. Therefore some subjects will have worked according to rules or have formulated their own rules, and others will have not. It is reasonable to assume that this might have had some effect on the outcome. Homer and Peterson (1980) suggested that DRO alone and DRO plus verbal instructions should be compared to assess the effectiveness of instructions; this study has not yet been carried out.

Telling clients the rules of a contingency may increase the speed and effectiveness of an intervention (Homer & Peterson, 1980; Poling & Ryan, 1982), and might make it easier to generalise the treatment to other settings. However, Hayes, Brownstein, Haas, and Greenway (1986) advocated caution in the use of rules. They pointed out that over-reliance on rules can make the person insensitive to changes in the contingency.

Studies which have compared momentary DRO and whole-interval DRO have shown differing results which could be interpreted in terms of the language abilities of the subjects. Repp et al. (1983) worked with children who had mild to moderate learning disabilities. They found whole-interval DRO to be a superior reductive treatment to momentary DRO, (although as outlined in previous chapters, they made use of highly explicit instructions which may have weakened the effect of momentary DRO). Derwas and Jones (1993), on the other hand, found momentary DRO to be as or more effective

than whole-interval DRO when used with people who had no language. Similarly, Sisson et al. (1988) found that momentary DRO alone was effective for just one of their three subjects; unlike the other two, that subject had no verbal skills.

Experiments 3 and 4 in the present thesis showed that momentary DRO can be effective with people who have language skills. Subjects in these experiments were provided with full instructions regarding the contingencies of momentary DRO and whole-interval DRO. The results showed that providing children with these instructions did not appear to inhibit the effectiveness of momentary DRO. However, the results are not conclusive. It is possible that the subjects may not have fully understood the rules of momentary DRO, or they may not have distinguished it from whole-interval DRO. In addition, it is possible that had instructions been less thorough, momentary DRO would have been even more effective than shown in these studies. It was proposed in Chapter 2 that in order to be fully effective, the rules of momentary DRO may need to be slightly ambiguous, otherwise capable subjects might exploit the contingencies and display the target behaviour at high rates during an interval.

EXPERIMENT 7

The effect of partial rules on the effectiveness of momentary DRO and whole-interval DRO

INTRODUCTION

This experiment investigated the effect of the two DRO schedules on non-learning disabled adults. This subject group was chosen for two reasons. Firstly, to see if both schedules were equally effective in reducing behaviour in this population. Secondly, because adults might be more adept at forming and using rules than children, and thus it would be possible to see more clearly the effect of language skills on the outcome of the schedules. It was decided not to provide subjects in this experiment with specific instructions regarding momentary DRO. This was in order to assess what happens when subjects are not told the difference between the two schedules, and whether any differences between the schedules would arise in such circumstances.

Studies which have compared whole-interval DRO and momentary DRO have all used single-case experimental designs. However, as no study has yet compared the two schedules in non-learning disabled adults, it was decided that a group design should be employed so that any major differences between subjects could be easily identified. The difficulty was to find a group of non-disabled subjects who shared a similar undesirable behaviour that would lend itself to controlled study. Although smoking was considered, subjects' smoking patterns would tend to vary greatly, and so make comparison more difficult. Finally it was decided to produce a simulated 'undesirable' behaviour so that subjects could all be assessed under similar conditions. It was important that this simulated behaviour should occur at a high rate, so that the effect of the schedules could be quickly seen. A computer game was chosen in which the subject was required to play 'tag' by

moving a symbol round the screen in pursuit of another symbol, operated by the computer. The simulated undesirable behaviour was designated as the subject moving their symbol into a particular area of the screen. This would occur at a high rate and subjects would all be equally likely to present this behaviour.

It was hypothesised that there would be a difference in the effectiveness in reducing the target 'undesirable' behaviour between whole-interval DRO and momentary DRO. It was also hypothesised that subjects who correctly guessed the designated area would show greater reduction in target behaviour than subjects who guessed the wrong area.

The aims of this experiment were as follows:
1. To determine whether there are differences between whole-interval DRO and momentary DRO in reducing a high-rate behaviour in non-learning disabled subjects.
2. To assess the effect of a partial rule on the performance of the two schedules.
3. To determine whether subjects who determined some more of the rule show different rates of behaviour to subjects who do not.
4. To provide further information regarding the different levels of reinforcement provided by the two schedules.

METHOD

Subjects

The subjects in this experiment were 30 volunteers recruited from a student subject pool. All subjects were undergraduate psychology students at University of Wales, Bangor. There were 17 females and 13 males. Subjects earned course credits for their participation, and they also earned small amounts of money in the form of reinforcers.

Apparatus and task

The task used in this experiment was a computer game called 'Tag'. This consisted of two symbols, labelled as 'A' and 'B' as in Figure 7.1.
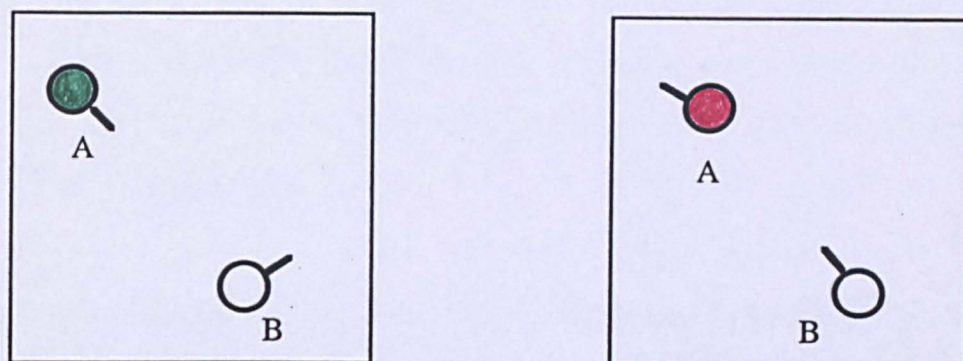
Figure 7.1 : Depiction of the 'Tag' computer game

The green symbol, 'A', was operated by the subject and the white symbol, 'B', was operated by the computer. 'A' could be moved forward in the direction of the pointer by holding down the 'Return' key on the computer. It could be rotated in a clockwise direction by pressing the 'X' key and anti-clockwise by pressing the 'Z' key. The object of the game was for the subject to move 'A' round the screen chasing 'B' (as shown in the box on the left). As soon as 'A' had caught 'B' (by touching the 'B' symbol), 'A' turned red and the positions were reversed (as shown in the box on the right). The 'B' symbol, operated by the computer, now began to chase 'A', operated by the subject. When 'B' caught 'A' the positions were again reversed with 'A' now chasing 'B'.

The computer which was used to display the game was an Acorn 3000 with a 14-inch screen. The screen was covered with a transparent acetate sheet, which was divided into six equal sections. These cross-sections were clearly visible to subjects, who were told that they would earn a small amount of money for keeping their 'Tag' symbol out of one of these sections for a certain length of time. Subjects were not told which section it was that they should avoid, nor how long they should avoid it for. The section of the screen which

was designated as the target area varied for each subject, randomly selected from the four outer areas of the screen. The two inner areas were not included as potential designated areas as a pilot study showed that subjects spent far greater amounts of time in the inner areas than the outer areas. Thus the behaviour designated as targeted for reduction by DRO schedule was entering a particular section of the screen. When the subject had their symbol in any other section of the screen than the designated one, this was regarded as an absence of the target behaviour.

For each subject, five minutes of baseline observation were taken. This was followed by fifteen minutes of intervention, during which reinforcers could be earned. The intervention was divided into three identical five-minute phases for the purposes of analysis, although observation and intervention carried straight on through the fifteen minutes without a break. Each of the four stages, baseline and phases 1, 2 and 3 of the intervention corresponded to a five-minute block on the recording sheet.

A ten-second momentary time-sampling procedure was used to record the occurrences of the target behaviour throughout the experiment. Recording sheets, a personal cassette-player and headphones, pre-recorded tape and stop-watch were used, as in Experiment 1.

Monetary reinforcers were provided, two pence for each successful fulfilment of the requirements of the DRO schedule, up to a maximum of £1.80 per subject. All subjects completed a brief questionnaire at the end of the experiment (see Appendix C).

Design

A between-subjects design was employed in this experiment, with subjects randomly allocated to one of two conditions: whole-interval DRO or momentary DRO.

There were three independent variables:

(i)   The type of schedule - whole-interval DRO or momentary DRO.

(ii)   Whether the subject correctly or incorrectly guessed the designated target area.

(iii)   Four experimental phases in which the behaviour was recorded: baseline, in which observation only took place, and phases 1, 2 and 3 in which reinforcement was delivered.

The dependent variable was the frequency of target behaviour.

Inter-response times

Prior to the experiment, a pilot study observed the responses of seven subjects who were asked to play the computer game for five minutes. The average length of time that they avoided entering a designated area of the screen was found to be 11.56 seconds. The DRO interval was set at 10 seconds for all subjects. It was felt that it would be difficult to calculate a new IRT for each subject during baseline, because there would be a gap between baseline and intervention while the IRT was calculated that would mean the subject having to wait. It would also make the experimenter's task more difficult during baseline if they had to time absences of the behaviour as well as take momentary time-sampling observations.

Procedure

Subjects were given written instructions at the beginning of the experiment. They were told that they were going to play a simple computer game for twenty minutes. It was explained that the first five minutes would be for practice only and that no points would be awarded. Subjects were told that during the following fifteen minutes of the game they would receive two pence for avoiding a particular area of the screen for a certain time. The aim was explained as being to receive as much money as possible (up to a maximum of £1.80) whilst playing the game to the best of their ability.

The rules of the computer game were then explained to subjects and they were shown the keys for moving their symbol round the screen. Any questions were answered at this point

and the baseline phase began when the experimenter was satisfied that the subject had understood the instructions.

Baseline

During this five-minute phase the experimenter recorded occurrences of the target behaviour using the momentary time-sampling procedure described in Experiment 1. No reinforcement was provided and there was no contact between experimenter and subject.

Whole-interval DRO

In this intervention subjects were given two pence for every complete interval of ten seconds that they avoided the designated area. The coins were placed in a small bowl which was on the table to the right of the subject, so that the subject could hear that they were receiving a reinforcer without having to look away from the screen. If a subject entered the designated area the timing ceased and the stop-watch was set to zero. A new interval did not begin until the subject had moved out of the designated area.

Momentary DRO

In this intervention subjects were given two pence if they were not in the designated area at the moment between two intervals. Occurrences of the target behaviour during the interval were irrelevant. The coins were placed in a bowl to the side of the subject, as for whole-interval DRO.

At the end of the experiment subjects were asked to complete a brief questionnaire, in which they were asked to indicate on a grid the area they thought they had been required to avoid.

Inter-observer agreement

For 50% of the sessions, observations were taken by a second observer to assess reliability of the recording method. Inter-observer agreement was calculated using the method

described in Chapter 1. The results showed a very high level of reliability as agreement averaged at 99.56%.

RESULTS

Percentage of time spent in designated area

Group comparisons

The mean percentage of time spent in the designated target area during the baseline phase and three phases of intervention can be seen in the Table below.

Table 7.1 Means and standard deviations of percentage of time spent in the designated area between phases and schedules

	Whole-interval DRO	Momentary DRO
Baseline	20.67　(SD = 9.69)	22.67　(SD = 6.92)
phase 1	17.33　(SD = 9.02)	19.11　(SD = 7.71)
phase 2	17.11　(SD = 10.90)	10.67　(SD = 9.36)
phase 3	15.56　(SD = 10.96)	11.56　(SD = 10.22)

At phase 2 the schedules diverged with much less time spent in the designated area in momentary DRO than in whole-interval DRO. This difference can be seen more clearly in Figure 7.2, where the downward slope in momentary DRO is much steeper than in whole-interval DRO.
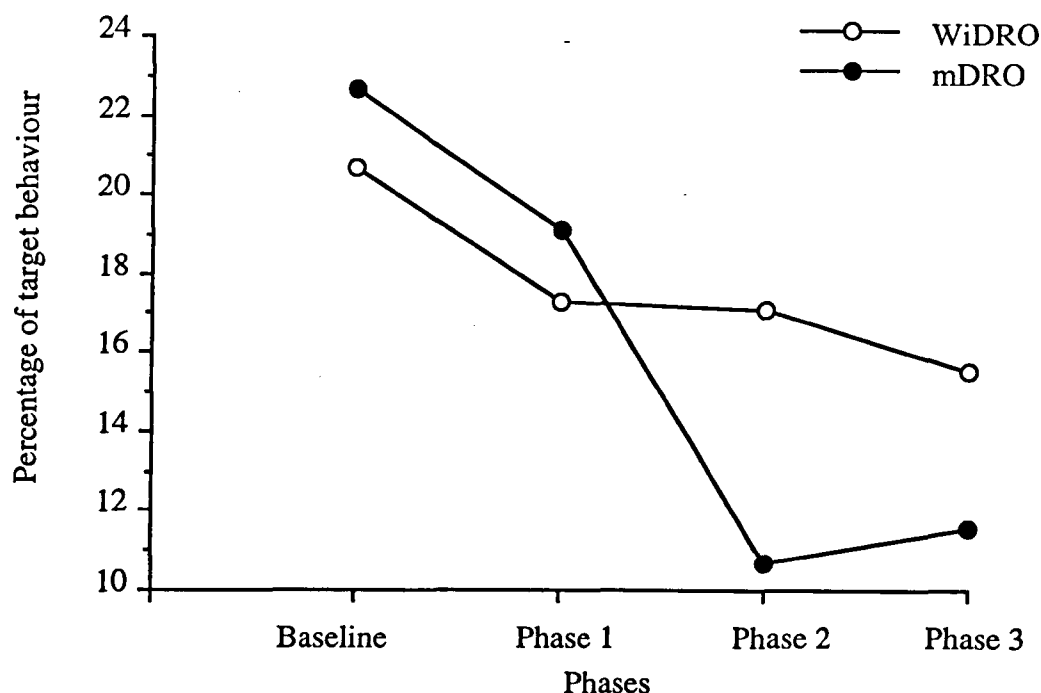
Figure 7.2 : The percentage of target behaviour in each phase of the experiment.

A two-way mixed analysis of variance was carried out to examine any difference across schedules within the various phases. There was one between-subjects factor, DRO schedule, which consisted of two levels, whole-interval DRO and momentary DRO; and one within-subjects factor, phase of experiment, which consisted of four levels - baseline, phase 1, phase 2 and phase 3. There was no main effect of schedule - that is, overall there was no difference in target behaviour between momentary DRO and whole-interval DRO. There was, however, a significant effect of phase ($F(3, 84) = 6.47$, $p<0.001$). This means that the target behaviour during some phases differed significantly from target behaviour in other phases. There was no significant interaction between schedule and phase, although the test did approach the 10% level of significance, suggesting a trend. This trend suggested the possibility of differential effects of phase in each schedule. Two further one-way ANOVAs were conducted to examine the effects of phase on performance in each schedule in more detail. The first ANOVA compared the four phases of the experiment in whole-interval DRO against each other. The second ANOVA compared the four phases in momentary DRO against each other.

The whole-interval DRO ANOVA showed no significant differences. In other words, the amount of time spent in the target area in any phase in whole-interval DRO did not differ significantly from the amount of time spent in it in any other phase. The momentary DRO ANOVA, however, did show a highly significant difference ($F_{(3, 56)} = 6.831$, $p < 0.001$). The table below shows the comparisons which were significant, according to Fishers PLSD post-hoc test.

Table 7.2 Significant phase comparisons within momentary DRO

Baseline vs phase 1	Not significant
Baseline vs phase 2	Significant
Baseline vs phase 3	Significant
Phase 1 vs phase 2	Significant
Phase 1 vs phase 3	Significant
Phase 2 vs phase 3	Not significant

This shows that there was a significant reduction in target behaviour in phase 2 from baseline, phase 3 from baseline, phase 2 from phase 1 and phase 3 from phase 1.

Individual comparisons

The mean of target behaviour was calculated for phases 1, 2 and 3 (e.g. all the intervention phases), thus forming one overall measure of behaviour during intervention. This measure was used to determine whether a reduction in target behaviour had occurred between baseline and intervention for each subject. The results are shown in Table 7.3.

Table 7.3 The number of subjects whose target behaviour was reduced by at least 50% during intervention and the number of subjects whose behaviour increased during intervention

	Whole-interval DRO	Momentary DRO
No. of subjects whose target behaviour was reduced during intervention by 50% or more from baseline	4	7
No. of subjects whose target behaviour increased from baseline during intervention	8	2

This table shows that momentary DRO produced more reductions in target behaviour than whole-interval DRO. Almost twice as many subjects in momentary DRO had their target behaviour reduced by 50% or more from baseline than subjects in whole-interval DRO. Over half the subjects in whole-interval DRO increased their target behaviour from baseline during intervention, compared to just two subjects in momentary DRO.

Effect of awareness on target behaviour

Subjects were asked after the experiment to identify the area that they thought they had to avoid in order to earn reinforcers (the target, or designated, area). It was hypothesised that subjects who correctly identified the designated area would spend less time in that area than subjects who incorrectly identified it.

The collapsed mean for phases 1, 2 and 3, as used above, was again employed as a measure of intervention. The mean amount of time spent in the target area during intervention by those who correctly and incorrectly identified that area is shown in the table overleaf.

The means show that in both schedules subjects spent less time in the target area if they correctly identified that area. Two independent t-tests show that these were significant differences. For whole-interval DRO, $t(13) = -1.77$ ($p<0.05$, one-tailed), for momentary DRO, $t(13) = -1.991$ ($p<0.05$, one-tailed).

Table 7.4 Amount of time spent in target area during intervention by those who did and did not identify that area

	Whole-interval DRO	Momentary DRO
Correct identification of target area	13.33 (N = 8)	11.55 (N = 10)
Incorrect identification of target area	20.48 (N = 7)	18.22 (N = 5)

Four repeated measures t-tests were carried out to explore the differences in target behaviour between baseline and intervention for subjects who did and did not correctly identify the designated area. (The collapsed mean for phases 1, 2 and 3 was again used as the measure for intervention.) The means are shown in the table below.

Table 7.5 Means of target behaviour in baseline and intervention across schedule and awareness of designated area

	Whole-interval DRO		Momentary DRO	
	Baseline	Intervention	Baseline	Intervention
Correct identification	23.75	13.33	22.67	11.55
Incorrect identification	17.14	20.48	22.67	18.22

The means show that in both schedules, target behaviour decreased during intervention amongst subjects who correctly identified the designated area. This difference did not quite reach statistical significance in whole-interval DRO, although it was highly significant in momentary DRO ($t(9) = 4.719$, $p<0.01$ two-tailed). Subjects who incorrectly identified the target area in momentary DRO showed a very slight reduction in target behaviour during intervention, though this was not significant. Performance from subjects who incorrectly identified the target area in whole-interval DRO was worse during intervention than baseline, although again this was not a significant difference.

The performance of subjects within each phase who were correct and incorrect within each phase was examined. The table below gives the means across schedules.

Table 7.6 Means of target behaviour across schedule and within phase for subjects who correctly and incorrectly identified the designated area

	Whole-interval DRO			Momentary DRO		
	Phase 1	*Phase 2*	*Phase 3*	*Phase 1*	*Phase 2*	*Phase 3*
Correct	15.83	15.00	9.17	18.33	8.33	8.00
Incorrect	19.05	19.52	22.86	20.66	15.33	18.67

A three-way mixed analysis of variance was used to assess differences across schedule and within phase, with regard to whether subjects had correctly identified the designated area. There were two between factors: (i) schedule, which consisted of two levels, whole-interval DRO and momentary DRO and (ii) identification of target area, which consisted of two levels, correct and incorrect. There was one within factor, phase of intervention, which consisted of three levels, phase 1, phase 2 and phase 3. Baseline phases were omitted from this analysis, as subjects' identification of the target area could not have had any bearing on their behaviour in baseline.

The three-way ANOVA showed that, as with the two-way mixed ANOVA described above, there was no effect of DRO schedule. However, consistent with previous tests, there was a significant effect of subjects' awareness of the designated area ($F_{(1, 26)} = 6.87$, $p<0.01$, one-tailed). That is, there was a difference in target behaviour between subjects who did and did not correctly identify the target area. There was no interaction between awareness and DRO schedule; similar patterns were seen amongst subjects who did and did not identify the target area in both schedules. There was an overall difference within phases, although this difference did not reach the 5% level of significance ($F_{(2, 52)} = 2.58$, $p<0.1$). That is, target behaviour tended to decline across phases. As before,

there was not a significant interaction between schedule and the various phases, although there was a trend.

In whole-interval DRO, subjects who correctly identified the designated area entered that area less than subjects who incorrectly identified it, but not considerably so, during phases 1 and 2. It was in phase 3 that subjects who correctly identified the area entered it far less. In momentary DRO, subjects who correctly identified the designated area entered it in phase 1 as much as subjects who incorrectly identified it. However, by phase 2 these subjects were entering the area much less frequently, and this reduction was maintained during phase 3. The differences can be seen in Figures 7.3 and 7.4 below.
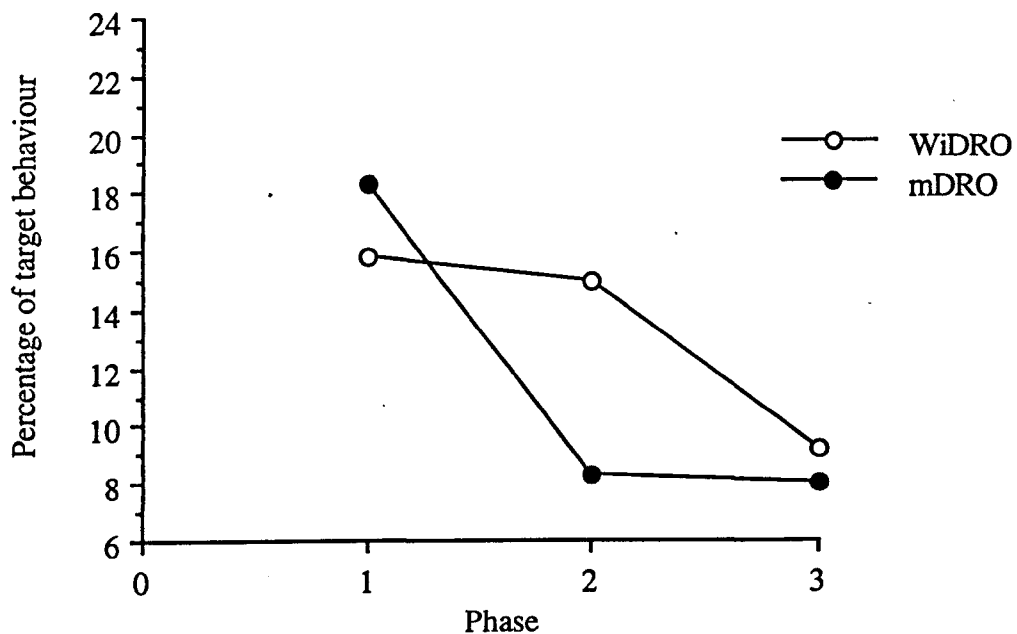
Figure 7.3 : The percentage of target behaviour in the three experimental phases by subjects who correctly identified the designated area.
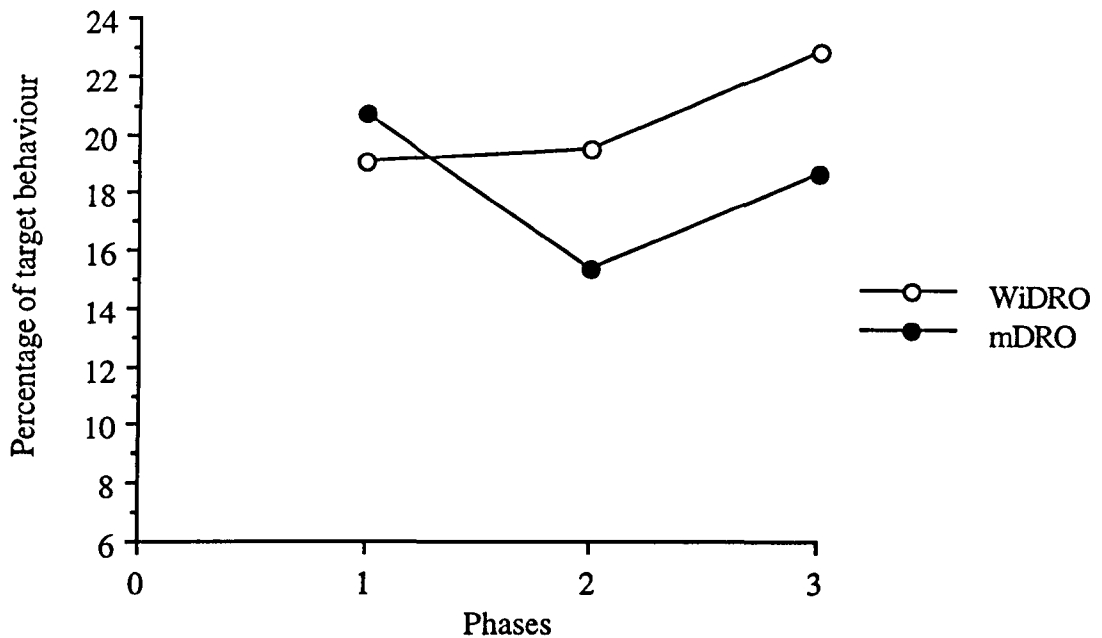
Figure 7.4 : The percentage of target behaviour in the three experimental phases by subjects who incorrectly identified the designated area.

There was an interaction between the various phases and subjects' awareness of the designated area, although this just missed the 5% level of significance ($F_{(2, 52)} = 3.00$, $p<0.1$). This means that a correct identification of the target area had a differential effect in the various phases: the later phases showed a lower percentage of target behaviour amongst those who correctly identified the area. This can be seen in Figure 7.5 in which momentary DRO and whole-interval DRO are combined.

There was not a three-way interaction between schedule, awareness and phase.
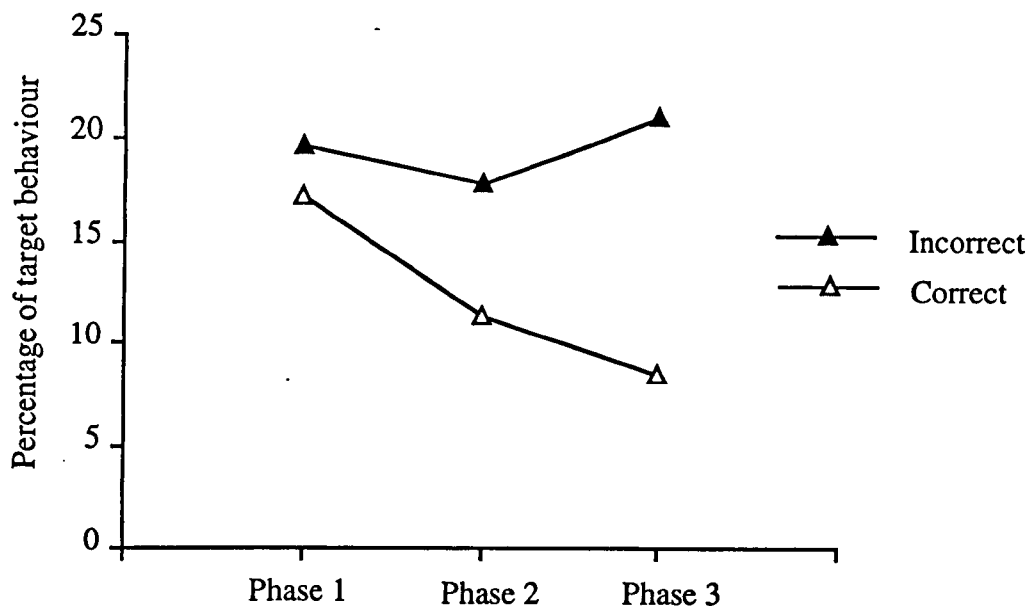
Figure 7.5 : The percentage of target behaviour in the three experimental phases by subjects in both schedules who correctly and incorrectly identified the designated area.

Amount of reinforcers earned

The maximum amount of reinforcers that could be earned in either schedule was ninety. The number of reinforcers earned overall by subjects in momentary DRO and whole-interval DRO were compared using an unpaired t-test. A highly significant difference was found between the schedules ($t(28) = -4.026$, $p<0.001$ two-tailed), with subjects in momentary DRO receiving considerably more reinforcers than subjects in whole-interval DRO. The means, standard deviations and ranges of the number of reinforcers can be seen in the table below.

Table 7.7 Means, standard deviations and ranges of number of reinforcers earned across schedules

	Whole-interval DRO	Momentary DRO
Mean	53.07	73.13
Standard deviation	16.718	9.65
Range	36 - 88	60 - 87

It can be seen from this table that there was much greater variability in the amount of reinforcers earned in whole-interval DRO than in momentary DRO, and that all subjects in momentary DRO earned at least two-thirds of the possible maximum number of reinforcers.

The amount of reinforcers earned by subjects who correctly identified the designated area was compared to the amount earned by subjects who did not identify the designated area. The means are depicted in the table below.

Table 7.8 Mean number of reinforcers earned across schedule by subjects who did and did not identify the correct target area

	Whole-interval DRO	Momentary DRO
Correct	60.88	76.50
Incorrect	44.14	66.40

It can be seen from this table that less reinforcers were earned when subjects incorrectly identified the target area. An unpaired t-test showed that this difference was significant in whole-interval DRO ($t(13) = 2.177$, $p<0.05$). There was also a difference in momentary DRO, although this test just missed the 5% level of significance ($t(13) = 2.142$, $p<0.1$). This means that subjects in whole-interval DRO who correctly identified the designated area earned significantly more reinforcers than subjects who did not identify it. There was a similar, though less pronounced, difference for subjects in momentary DRO.

DISCUSSION

To sum up the findings of this experiment: there was greater reduction of the target behaviour between phases in momentary DRO than whole-interval DRO. Momentary DRO reduced the target behaviour by at least 50% in more subjects than did whole-interval DRO.

Subjects in both schedules spent less time overall during intervention in the designated area if they correctly identified that area. The overall reduction in target behaviour from baseline to intervention was more marked amongst those who identified the correct area in momentary DRO than in whole-interval DRO. The phase in which correct identification first impacted upon performance was the second phase in momentary DRO and the third phase in whole-interval DRO. These findings are considered in more detail below.

Percentage of time spent in designated area

Group comparisons

There was no difference between subjects at the start of the experiment: subjects in both schedules spent similar amounts of time in the target area in baseline (around 20%). In the first intervention phase, there was again very little difference between the two schedules. Subjects in both groups spent approximately 3% less time in the target area during this phase than in baseline. It was in the second five minutes of intervention that the groups diverged. Subjects in momentary DRO spent less than half the amount of time in the target area in phase 2 than in baseline. Subjects in whole-interval DRO on the other hand, spent almost exactly the same amount of time in phase 2 than phase 1, neither of which differed significantly from baseline. In the third and final phase of intervention, subjects in momentary DRO showed a very slight rise in the amount of time spent in the target area, though this did not differ significantly from the amount of time spent in phase 2. Subjects in whole-interval DRO showed another small reduction, but again, one which did not differ significantly from any of the previous phases, including baseline.

Overall, then, momentary DRO showed a sharp and significant reduction between the first ten minutes (baseline and phase 1) and the second ten minutes (phases 2 and 3), while whole-interval DRO showed a very slight reduction from one phase to the next, none of which differed significantly either from its predecessor or from baseline. It is important to note that although momentary DRO showed a significant reduction in target behaviour and whole-interval DRO did not, there was not a significant difference in the amount of target

behaviour between the two schedules. That is, momentary DRO did not in any phase -
produce significantly lower rates of target behaviour than whole-interval DRO. This is
presumably because the two schedules were producing very similar results in baseline and
phase 1, and then, when momentary DRO showed a sharp reduction, target behaviour in
whole-interval DRO also fell, albeit less steeply.

It is possible that had the experiment continued, target behaviour in whole-interval DRO
would have continued to reduce slowly and steadily until it reached the low levels of
momentary DRO. This cannot be ruled out by the data presented here. What can be said
with certainty, however, is that momentary DRO effected a far speedier reduction in target
behaviour than whole-interval DRO.

Individual comparisons

Momentary DRO produced more individual reductions of target behaviour of 50% or more
than whole-interval DRO (7 and 4 respectively). More strikingly, whole-interval DRO
showed overall increases in target behaviour during intervention in more than half the
subjects. The same effect was seen in momentary DRO for only 13% of subjects. It is
possible that these results were related to the findings on awareness of the designated area
considered below. Whatever the cause, however, it is clear that momentary DRO produced
greater reductions than whole-interval DRO.

Effects of awareness on target behaviour

Subjects were asked if they could identify the area that they thought they had had to avoid
in order to earn reinforcers. It was originally thought there might be difficulties with
asking subjects to indicate the area after the experiment, in that they might choose the right
area even though they hadn't been avoiding it during the testing. But the results showed a
clear difference between those who did and those who did not select the right area,
suggesting that the area that subjects pointed out after the experiment tended to be the one
that they were avoiding during the experiment.

The results show a clear difference between subjects who did and did not correctly identify the designated area, regardless of schedule. There were also clear differences seen in the reduction from baseline to intervention between those who did and did not identify the correct area. A similar reduction was seen in both schedules, and this reduction reached statistical significance for momentary DRO. Presumably the reason why whole-interval DRO did not show a statistical difference was because the variances were comparatively high in both baseline and intervention (much higher than in momentary DRO).

Comparisons across phases between those who were correct and those who were not can be seen most clearly in Figures 7.3 and 7.4. Figure 7.3 depicts subjects who correctly identified the target area. The disparity between schedules in phase 2 can be clearly seen. By phase 3 the two schedules were almost identical. Figure 7.4 depicts subjects who incorrectly identified the target area. Although momentary DRO was lower than whole-interval DRO in phases 2 and 3, neither schedule showed a declining pattern as seen in Figure 7.3. Rather, both schedules rose at the third phase, confirming that these subjects were quite unaware of the correct area.

For subjects who were correct, the patterns for momentary DRO and whole-interval DRO were almost identical to the pattern seen in the overall phase graph (Figure 7.2). As with Figure 7.3, here again the difference between schedules occurred at phase 2, with momentary DRO falling away from whole-interval DRO. The difference between Figure 7.2 and Figure 7.3 is that in Figure 7.3 at phase 3, the two schedules converged again and were almost identical. This was accomplished by a sharp reduction in behaviour in whole-interval DRO; momentary DRO in phase 3 was almost exactly the same as in phase 2. It would be interesting to find out what would have happened to 'correct' subjects if the experiment had continued. It can be seen from Figure 7.3 that at phase 3, momentary DRO appeared to have levelled out, whilst the trend in whole-interval DRO was downwards. It is possible to speculate that those trends would continue: further experimentation would be necessary to clarify this.

Figure 7.4, depicting subjects who were incorrect, shows a different picture. Although momentary DRO and whole-interval DRO again diverged at phase 2, the difference was much smaller; and both rose in phase 3, although the difference between them remained the same as in phase 2. It appears then, that if subjects were operating under a false rule - namely, that the designated area was a segment other than the correct one - the amount of target behaviour rarely differed greatly from baseline. Although not significant, the consistent difference between the schedules in phases 2 and 3 (4.19% in both cases) suggests that even when subjects were wrong, momentary DRO resulted in lower levels of target behaviour than whole-interval DRO.

The individual data showed that in whole-interval DRO, only four subjects achieved reductions from baseline of 50% or more. All four subjects were people who correctly identified the designated area, suggesting that correct identification was a prerequisite for a substantial reduction in whole-interval DRO. It was not a sufficient factor, however - two subjects who identified the correct area showed an increase in target behaviour from baseline. Of the seven subjects in momentary DRO who achieved reductions of 50% or more, all but one identified the correct area. Again, this was not a sufficient factor: one of the two subjects whose behaviour increased identified the correct area. Therefore it seems that there was a relationship between individual performance and awareness of the target area, but this was not the only factor which contributed to reductions in target behaviour.

To sum up, the group results show that when subjects knew the rule, they ultimately did as well in whole-interval DRO as momentary DRO; however, the reduction was quicker in momentary DRO. When subjects did not know the rule, a reduction was not brought about by either schedule; but target behaviour was consistently lower in momentary DRO than whole-interval DRO.

Caution needs to be exercised with all the findings which are concerned with subjects who were incorrect in momentary DRO, as the numbers were relatively small. It is worth noting

that whilst more or less equal numbers of subjects in whole-interval DRO were correct and incorrect regarding the target area (8 and 7 respectively), in momentary DRO twice the numbers of subjects were correct (10) than incorrect (5). A larger replication of the present experiment could ascertain whether it is in fact easier to work out rules in momentary DRO than in whole-interval DRO. As outlined in Chapter 3, it is possible that this is a function of the increased reinforcement seen in momentary DRO: the greater number of reinforcers provided more 'clues' about which was the correct area and thus speeded up the learning process. This and other possible reasons for the effectiveness of momentary DRO are considered in more detail in Chapter 7.

What is not known is if the actual contingencies of momentary DRO were also easier to determine. Knowledge about the area itself may have been more accessible, but whether subjects realised they only had to avoid it at certain moments is doubtful. Further experiments would be needed to clarify this, but the fact is that the subjects in momentary DRO who correctly identified the target area had the lowest overall rate of target behaviour in the experiment. This suggests that although they identified the area they did not realise that they did not have to avoid it all the time. This is an important point which should be examined further by future experiments. It is still not clear from this experiment whether momentary DRO could work effectively if subjects definitely understood that they only had to refrain from the target behaviour occasionally. It is possible that in order for momentary DRO to work, subjects must be able to easily understand that they can earn reinforcement if they refrain from a particular behaviour, but be unable to work out that they don't have to refrain all the time. Further experiments could perhaps establish whether it is equally difficult - or easy - to work out the contingencies of momentary DRO and whole-interval DRO.

The findings outlined here suggest that if DRO schedules are used with people who are capable of forming rules, then momentary DRO should be the schedule of choice, as it

produced a quicker reduction, and reduced behaviour in more subjects than whole-interval DRO.

Number of reinforcers

As seen in previous experiments, much greater reinforcement was awarded in momentary DRO than in whole-interval DRO. The possible total amount of reinforcement was the same for both schedules (90), and the highest amount earned by individual subjects was almost identical (87 in momentary DRO and 88 in whole-interval DRO). However, the range was far greater in whole-interval DRO, with 47% of subjects earning less than half the possible total. All subjects in momentary DRO earned at least 60 reinforcers (two-thirds of the total possible), and 27% earned more than 80 reinforcers.

Unsurprisingly, subjects who identified the correct area earned more reinforcers than subjects who did not, in both schedules. However, the difference between 'correct' and 'incorrect' was more marked in whole-interval DRO than momentary DRO. This shows that subjects in momentary DRO could earn reinforcers even when they weren't avoiding the correct area, presumably because even if they weren't deliberately avoiding it they only had to remain outside it for brief moments and thus could still earn reinforcers whilst entering it. On the other hand, subjects in whole-interval DRO who were incorrect could not earn as many reinforcers as those who were correct, because even if they went into the target area just briefly, they lost the chance of reinforcement for that interval. This accounts for the more marked difference in reinforcement between correct and incorrect subjects seen in whole-interval DRO.

The means show that overall, subjects who were correct in whole-interval DRO nonetheless earned on average less reinforcers than subjects in momentary DRO who were incorrect. This appears to be simply a result of the different rates of reinforcement that can be earned in the two schedules, due to the nature of the schedules rather than any particular aspect of the experiment itself.

Conclusions

This experiment revealed a large difference in behaviour between subjects who did and did not correctly guess the target area. Essentially, both schedules were effective if subjects guessed the target area; and they were both ineffective if subjects did not guess the target area. This has important implications for using schedules with verbal subjects; if subjects know which behaviour is being targeted, both schedules are effective. It is important to note that the use of the wrong rule, which was shown in this experiment to be related to the ineffectiveness of the schedules, is not the same as the non-use of rules. The behaviour of subjects who do not use rules at all can still be reduced under DRO schedules.

Further evidence for the importance of the greater levels of reinforcement in momentary DRO was provided by this experiment. Subjects were given more information in the form of reinforcement regarding whether or not they were in the target area, and it is likely that this contributed to the speedier reduction in momentary DRO. Whether there is more to the efficacy of momentary DRO than density of reinforcement is considered in Chapter 7.

The possibility that momentary DRO would not work with highly verbal subjects was not supported by the evidence. If anything, momentary DRO was more effective than whole-interval DRO. This adds further weight to the suggestion that the reason why Repp et al. (1983) failed to show a reduction with momentary DRO was the use of very specific instructions, rather than the verbal ability of the subjects. This experiment, like Experiments 3 and 4, suggests that if momentary DRO is to be used with verbal subjects, then slightly ambiguous rules should be given to ensure maximum effectiveness. The optimum amount of information to be included in a rule could be determined experimentally, by comparing the effects of providing the correct rule with providing a partial rule and providing no rule. Such an experiment could be carried out with children with disruptive behaviour who could make use of rules. Differences between those who were given the complete rule and those who generated some or all of their own rule would provide some useful information about the nature of DRO schedules.

Overall, this experiment showed that momentary DRO effected a quicker reduction in target behaviour than whole-interval DRO in non-disabled adults. This quicker reduction could be largely accounted for by subjects who correctly identified the target area in momentary DRO. Subjects in whole-interval DRO who were correct took longer to achieve the same reduction; and subjects in whole-interval DRO who were incorrect performed consistently poorer than such subjects in momentary DRO.

Chapter 5

Momentary DRO as part of a treatment package

Whole-interval DRO has been noted for its compatibility with other treatments (Poling &

Ryan, 1982). In situations where DRO alone has not been highly effective, or where other

treatment goals are being concurrently targeted, whole-interval DRO has shown itself to be

easily combined with other treatments such as verbal reprimands, time-out and

overcorrection. It would be useful to find out if momentary DRO offered the same

flexibility.

EXPERIMENT 8

Discrimination training using a treatment package of momentary DRO, instructions, modelling and reprimand

INTRODUCTION

This experiment explored the use of momentary DRO as part of a treatment package. Momentary DRO was combined with modelling, instructions and reprimand in an attempt to apply discrimination training to stereotyped behaviour. The subject was taught, via the treatment package, that displaying stereotypy was acceptable in one situation and not in another.

An important applied question in dealing with inappropriate behaviours is whether attempts should be made to totally eliminate them, or whether they should be selectively reduced. For example, the DRL schedule has been used to reduce behaviours such as excessive talking and eating (Lennox, Miltenberger, & Donnelly, 1987) - behaviours which it would not be appropriate to extinguish completely. Another technique of selective reduction is discrimination training, though this has been little used with inappropriate behaviours (exceptions include Woods, 1983, discussed below).

Philosophical considerations

Several researchers have proposed that stereotyped behaviour should be selectively reduced. For example, LaGrow and Repp (1984) recommended that people should be taught to "...discriminate between free periods, learning periods, and social periods, with stereotypy considered inappropriate only in the latter two" (p. 607). Berg and Wacker (1991) also suggested that it would be more beneficial to use DRO schedules in selected situations than attempt to suppress stereotypy entirely. Proponents of this idea point out

that stereotyped behaviour is rarely physically harmful and that it need not interfere with adaptive learning (Klier & Harris, 1977; Wolery, Kirk, & Gast, 1985). Indeed, for some individuals, their stereotyped behaviour is one of the very few self-initiated responses they have. The main problem with harmless inappropriate behaviours such as stereotypy is that they are socially unacceptable (Jones, 1991b; Wolfensberger & Thomas, 1982) and may prevent an individual's integration into the community (Jones, Wint, & Ellis, 1990).

Discrimination training

An experiment carried out by Woods (1983) provides one of the few examples of an attempt to selectively reduce stereotypy using discrimination training. He taught a ten year-old girl who continuously leafed through magazines that it was acceptable for her to use magazines which had a large red triangle on the cover (S^d) and unacceptable for her to use those which did not (SΔ). A combination of brief immobilisation and praise was used to strengthen the appropriate responses, and Woods showed that discrimination between the two stimuli was achieved.

The present experiment also aimed to train a person with learning disabilities that a behaviour is acceptable in some circumstances and unacceptable in others. Woods acknowledged that his experiment was an analogue study rather than representative of a real situation because of the unnatural materials employed (acceptability determined by a red triangle would obviously not be encountered in the 'real world'). The present experiment, therefore, aimed to remedy this limitation by employing a more naturalistic discrimination - two different rooms. It is feasible that it might be considered socially acceptable for an individual to display stereotypy in one room, for example, the bedroom, but unacceptable in another, such as the dining room.

The present experiment combined momentary DRO with a number of other components in order to train the subject to discriminate between two stimuli. Interestingly, the subject in

this experiment displayed the same stereotypy as the subject in Woods' experiment: magazine flipping.

The aims of this experiment were:

1.   To investigate the utility of momentary DRO as part of a treatment package.

2.   To train a subject with learning disabilities to discriminate between two stimuli.

METHOD

Subject

Saul was 26 years old and had severe learning disabilities. He attended a day-centre in Oxford during the week. He had no self-help skills and did not have language, although he responded to simple commands and instructions.

Response definition

Saul had a small cataract on each of his eyes which meant that his eye-sight was fairly restricted. This seemed to have some bearing on his only stereotyped behaviour, which was magazine flipping. This was defined as holding a magazine close to his face and turning the pages of a magazine rapidly in front of his eyes. Because Saul did not pick up a magazine unless to display the stereotypy, the definition of the behaviour was broadened to include holding a magazine as well as flipping it. The type of magazines Saul preferred were brightly coloured holiday brochures. Saul's key-worker completed the Motivation Assessment Scale and this suggested that the stereotypy was maintained by sensory stimulation.

Apparatus and discriminative stimuli

A cassette-player and headphones, pre-recorded audio-tape, recording sheet and stop-watch were used, as described in Experiment 1 (Chapter 2). Two rooms were used as the discriminative stimuli, Room 1 and Room 2, which were both very different in furnishing

and lay-out. Room 1 (SΔ) was designated as the room where Saul was not permitted to display stereotypy. This room was an office which contained desks, chairs, a sink and a coffee machine. Room 2 (S^d) was designated as the room where Saul was permitted to display his stereotypy. The walls, carpet and furniture in this room were all blue. A pile of holiday brochures was placed in each room next to the chair set aside for Saul. A cassette-player and a music tape were placed in Room 1 to provide reinforcement. Access to Room 2 was restricted and thus in the intervention phase and follow-up there were slightly fewer sessions in Room 2 than Room 1.

Design

A multiple-schedule ABA AC design was employed in this study. Phases A, B and A formed the main study and phases A and C formed the follow-up study. The multiple-schedule design was chosen as it allows each intervention to be paired with a particular stimulus. Baseline measures (A) were taken in both rooms, so that any effect of the different setting on behaviour could be noted.

Phase B consisted of alternating treatment sessions. The treatment in Room 1 (SΔ) consisted of four components, all designed to show the subject that magazine flipping was not acceptable in this room. The four components were:

(i)   Momentary DRO, which was used to reinforce the absence of the target behaviour.

(ii)   Reprimand of subject, which was used to punish the presence of the target behaviour.

(iii)   Modelling, shown by Bandura (1965) to be a powerful learning technique.

(iv)   Reprimand of model. Bandura showed that models who were punished for their behaviour were considerably less likely to be imitated by children than models who were not punished.

The treatment in Room 2 (S^d) consisted of continual modelling without reprimand, which was designed to show the subject that magazine flipping was acceptable in this room. This

type of imitation has been shown to be reinforcing (e.g. Miller & Morris, 1974), and also indicates to the subject that such behaviour is permissible. On each day, one treatment took place at 9.00am and the other at 11.00am. The treatments were counter-balanced across times to avoid the possibility of time of day or order effects. Prior to each treatment, Saul was given instructions regarding whether or not he was allowed to play with his magazines in that particular room.

The second phase A was a return to baseline, which was used to determine if any lasting discrimination had been formed during treatment which remained when treatment was removed.

This part of the experiment took place over three weeks.

Ten weeks later the follow-up study took place. This consisted of phase A (Baseline) and phase C (momentary DRO alone). This follow-up was a partial component analysis which aimed to determine the contribution of momentary DRO to the overall treatment package. This part of the experiment took place over two days. Baseline measures were taken on the first day, with Room 2 sessions taken at 9.00am and Room 1 sessions taken at 3.00pm. On the second day treatment sessions were carried out, with Room 1 sessions taken at 9.00am and Room 2 sessions taken at 3.00pm.

Session length during all phases of the experiment was five minutes.

Reinforcer

The results of the Motivation Assessment Scale suggested that Saul's stereotypy was maintained by sensory factors. Several members of the care-staff suggested that Saul responded very positively to music, and so this was the reinforcer chosen for this study.

Inter-response time

The length of time between occurrences of Saul's magazine flipping was recorded over two days. The average length between occurrences was 29.14 seconds. The DRO interval length was set at 20 seconds.

Procedure for the main study

Baseline (A)

Baseline measures were taken in both rooms. At the beginning of each block of sessions, Saul and the two experimenters walked together down the corridor to where both rooms were situated. During this period nothing was said to Saul. The same procedure was followed on entering both rooms. Saul was asked to sit down and a pile of magazines were placed at his feet. Both experimenters then sat down. Occasionally the second experimenter would tidy the magazines into a pile as Saul tended to discard magazines around his chair. Observations of the stereotypy were taken. No other contact was made between Saul and the experimenters, nor between the experimenters.

Intervention (B)

During this phase, Room 1 was designated as an unacceptable room for Saul to display the stereotypy, and Room 2 was designated as an acceptable room for this behaviour.

Room 1 (SΔ)

When Saul was walking to this room at the start of the block of sessions, the first experimenter explained that they were going to Room 1, and reminded him that in this room he was not allowed to play with his magazines. If Saul was holding a magazine it was taken away at this point. Once inside the room, Saul sat down on the chair that had a pile of magazines by it, and the first experimenter sat opposite him. The second experimenter sat by the cassette-player which was situated in the corner of the room.

The first experimenter recorded observations of the behaviour, operated the momentary DRO schedule, and delivered reprimands. The second experimenter modelled Saul's stereotypy every ten minutes, and operated the cassette-player. The four components to this intervention are described below.

(i) Momentary DRO and (ii) Reprimand of subject

Saul was given a reinforcer if he was not holding or flipping a magazine at the moment between two 20-second intervals. The first experimenter signalled to the second experimenter that Saul had earned a reinforcer, and the second experimenter played ten seconds of music on the cassette-player. Timing of the DRO interval was suspended while the music was playing. If Saul was holding or flipping his magazine at the moment between two intervals, the first experimenter reprimanded him (by saying "No") and removed his magazine and placed it on the floor. No reinforcer was given, and another 20 second interval then began.

(iii) Modelling and (iv) Reprimand of model

The second experimenter modelled Saul's stereotypy once every ten minutes. The first experimenter then turned to him and said "No, you're not allowed to do that in this room". The second experimenter then put the magazine down on the floor.

Twenty-four sessions of this treatment package were carried out.

Room 2 (S^d)

When Saul was walking to this room at the start of each block of sessions, the first experimenter explained that they were going to Room 2, and reminded him that in this room he was allowed to play with his magazines. Once inside this room, Saul sat on the chair which had magazines by it, and the first experimenter sat down opposite him and took observations of the behaviour. The second experimenter sat next to Saul and began modelling the stereotypy. This modelling carried on throughout the session. The

experimenters interacted occasionally, by talking about matters unrelated to the magazine flipping. This was to show Saul that the behaviour was acceptable in this room and that the first experimenter was not reprimanding the second experimenter for this behaviour.

Twenty-one sessions of this intervention were carried out.

Return to baseline (A)

This phase was the same as the initial baseline. Observations were taken in both rooms, as before.

Procedure for the follow-up study

The follow-up was designed to examine the role of momentary DRO within the treatment package. It also aimed to investigate whether the effects of the first study had lasted during the period between main study and follow-up.

Baseline (A)

This baseline phase was identical to the previous two baseline phases.

Intervention (C)

Room 1 (SΔ)

Saul was given no verbal instructions on the way to the room. He sat down in the same chair as in the main study which again had magazines by it. The momentary DRO procedure was carried out as described above. However, if Saul was flipping a magazine at the moment between two intervals, he was not reprimanded and the magazine was not taken away from him. Modelling and reprimand of model did not take place in this treatment.

Room 2 (S^d)

Saul was given no instructions on the way to the room. Modelling by the second experimenter was carried out during these sessions, but there was no contact between the experimenters.

Inter-observer agreement

Observations were taken by a second observer during follow-up only. They were taken for all sessions during this part of the study. Agreement between raters was calculated as for Experiment 1, and it ranged from 97% to 100% with a mean of 99%.

RESULTS

Figure 8.1 shows the percentage of stereotypy which occurred during all phases of this study.

Main study

During the initial baseline phase magazine flipping in Room 1 ranged from 47% to 93%, averaging 73%. In Room 2 it ranged from 70% to 100%, averaging 86%. During the intervention phase magazine flipping in Room 1 ranged from 0% to 20%, averaging 6%. In Room 2 it ranged from 50% to 100%, averaging 88%. During the return to baseline phase magazine flipping in Room 1 ranged from 0% to 100%, averaging 46%. In Room 2 it ranged from 77% to 100%, averaging 97%.

Follow-up

During baseline in the follow-up study, magazine flipping in Room 1 ranged from 63% to 100%, averaging 86%. In Room 2 it ranged from 97% to 100%, averaging 99%. During intervention magazine flipping in Room 1 ranged from 87% to 100%, averaging 96%. In Room 2 it ranged from 60% to 93%, averaging 81%.
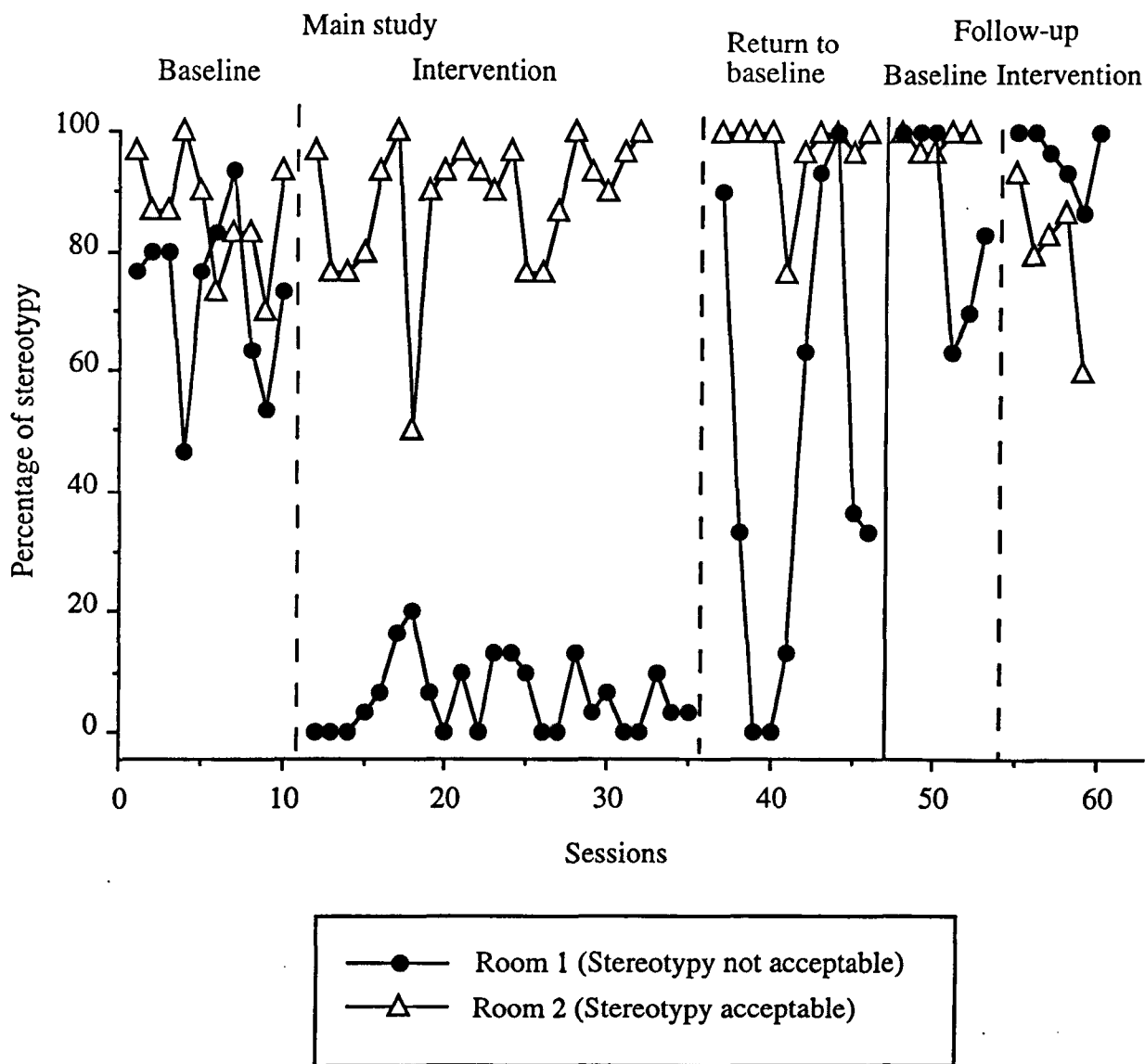
Figure 8.1 : The percentage of Saul's stereotypy which occurred during all phases of the experiment.

DISCUSSION

The results of this study show that the package of treatments used in Room 1 in the (B) phase were highly successful in reducing Saul's stereotyped behaviour. The follow-up study shows that momentary DRO alone did not achieve a reduction. These results will be considered in more detail.

142

During baseline there was a small difference in the rate of Saul's stereotypy in the two rooms. Although there was some overlap, 80% of sessions in Room 1 were lower than those in Room 2. There were no differences between procedures in the two rooms in this phase; the most likely explanation for this difference is that Saul found some of the items in Room 1 distracting and thus displayed less stereotypy in that setting. In particular, there was a bookcase that Saul approached and examined closely at the start of each session. This was ultimately covered over by a rug so that it no longer provided a distraction. In contrast there was almost no furniture in Room 2, and few objects to look at. Although the difference between baselines in the two rooms is not large, it must be borne in mind when considering the results of the intervention. Nonetheless, the data from this baseline phase show that Saul's magazine flipping tended to occur at a very high rate when he was not engaged in other activities.

In phase B there was an immediate and dramatic decrease in stereotypy in Room 1. Magazine flipping fell to zero as soon as the treatment package was introduced, and remained low throughout the remainder of the phase. The fact that there was such an immediate reduction in Room 1 suggests that the initial instructions which were given to Saul as he walked to this room were effective. Because Saul's stereotypy was low right from the start of the intervention, it was possible for him to earn a high level of reinforcement and this may in turn have kept the stereotypy at a low level. During the fourth session Saul began to flip a magazine and received his first reprimand. At first, he seemed confused by this. A short period began where he would slowly reach for a magazine and start flipping it, until reprimanded again. It appeared that he was checking to see if he would be told not to flip the magazine each time. After he had been reprimanded several times, he reached for the magazines less frequently.

The second experimenter had been reprimanded for magazine flipping several times before Saul received his first reprimand. However, until this point Saul had taken little notice of this. After he himself had been reprimanded, however, he began to pay more attention.

Often, Saul laughed when the experimenter was reprimanded and his magazine was taken away, particularly if Saul was still holding his. It appears that some kind of connection between himself and the second experimenter was acknowledged by Saul. He seemed to 'side' with the second experimenter against the first experimenter and often grinned at him and made eye contact, especially when the second experimenter was reprimanded.

The role of momentary DRO in the treatment package was not clear, as Saul appeared to play little attention to the music when it was played. However, on several occasions when the music was not playing he put down a magazine, went over to the cassette-player and started touching it. This suggests that he thought the music should be playing immediately he put down his magazine, which suggests that he was learning the relationship between the absence of stereotypy and the music.

In contrast to Room 1, stereotypy in Room 2 remained high, with a mean percentage of behaviour similar to that in baseline. Saul again appeared to demonstrate a rapport with the second experimenter who modelled his behaviour. Often Saul would watch the model from behind his magazine and on several occasions chose a magazine that the model had used. Occasionally he would smile if he made eye-contact with the model. It appeared that Saul took his cue from the model that it was acceptable to flip magazines in the room.

Several other factors demonstrated to Saul that he was allowed to play with his magazines in this room. These included the instructions on the way to the room, and the lack of reprimands. In addition, Saul was not reinforced for the absence of the stereotypy, as was the case in the other room. All these undoubtedly contributed to showing Saul that his behaviour was acceptable in this room; but the modelling seemed to perform a particularly important role. It is interesting to note that despite Saul being permitted, indeed encouraged, to display his stereotypy, the behaviour did not increase from baseline. This indicates that a selective discrimination need not produce higher levels of stereotypy in the setting where this behaviour is deemed acceptable.

By the end of the intervention phase, Saul appeared to have formed a discrimination between the two rooms. As he entered Room 2 he invariably reached for a magazine, whereas this rarely happened in Room 1.

During the return to baseline phase, all aspects of the interventions were removed. There was little change in stereotypy in Room 2: it was very slightly higher than during initial baseline and intervention phases. However, there was an erratic pattern of behaviour in Room 1. In the first session stereotypy rose to 90%. This may have occurred in part because this was the first session after the intervention in which instructions were not given to Saul as he walked to the room. This suggests again that the pre-session instructions used in the treatment phase were important in producing an initial reduction. Once Saul entered the room he appeared to wait for a few minutes, perhaps for the music to be played. He then slowly picked up a magazine, all the while looking at the first experimenter. He then began to flip the magazine rapidly and, in the absence of any reprimand, continued to do this throughout most of the session. Although there were some sessions where low amounts of stereotypy occurred, the pattern during this phase suggests that once the treatment package was removed, the effect of the learning began to decrease. The trend during this phase in Room 1 suggests that the behaviour would slowly stabilise, with the falls in the rate of behaviour gradually becoming less steep, until the behaviour reverted to a pattern similar to that in baseline.

This prediction is supported by the baseline phase in the follow-up study, which showed that generalisation over time had not occurred. Both rooms contained levels of stereotypy higher than in the initial baselines. The pattern of results in the baseline phase in follow-up was similar to that in the initial baselines, as behaviour in Room 2 was slightly higher than that of Room 1. The last three sessions in Room 1, although on an upward trend, were much lower than the first three sessions of 100%. It may be that, as in the initial baseline, Saul found more to distract him in this room than in Room 2, though it is unlikely that this should have occurred during the latter sessions only. The pattern of results is more likely

to be due to a time of day effect. The time of day was not a problem in the main study, because both treatment sessions took place in the morning. However, in the follow-up study, baseline sessions in Room 2 were taken in the morning (9.00am) and the sessions in Room 1 were taken in the late afternoon (at 3.00pm), when Saul was preparing to go home and was more tired and restless. During these afternoon sessions Saul spent more time just sitting still or wandering round the room and less time playing with the magazines. This suggestion is supported by the data in the follow-up intervention phase, where the sessions were reversed so that the data in Room 2 were taken in the afternoon. They followed a similar pattern to the Room 1 baseline sessions in that they were generally lower than the sessions in the morning. In fact, this was the only phase in the entire experiment in which the stereotypy in Room 2 was lower than in Room 1. This suggests that Saul's stereotypy was affected by tiredness, and that the intervention phase in the follow-up was effected by the time of day.

The results show that momentary DRO alone had no effect on Saul's stereotypy. Indeed, the stereotypy in Room 1 during this phase was higher than in any other phase, suggesting that momentary DRO alone may even have increased the stereotypy. However, what actually happened was that momentary DRO did not get a chance to work. Initially, the absence of instructions on the way to the room meant that as soon as he entered the room, Saul picked up a magazine. Secondly, the absence of reprimands when Saul was magazine flipping meant that he rarely stopped and thus reinforcement for the absence of stereotypy could not be delivered. Indeed, a reinforcer was awarded only three times during the six sessions of this phase. If the stereotypy had occurred less frequently, then reinforcement for its absence could have been given, and the effectiveness or otherwise of momentary DRO could have been seen. Therefore, a statement about the contribution of momentary DRO to the treatment package cannot be made, as it had so little chance to operate during this phase. This suggests that with very high-rate behaviours such as Saul's, a combination of instructions and reprimand to stall the behaviour and momentary DRO to reinforce its absence may be more effective than one component alone.

Conclusions

This experiment showed that a person with learning disabilities could discriminate between two stimuli and learn that stereotypy was acceptable in the presence of one stimuli and unacceptable in the presence of the other. Woods' (1983) findings were extended by this present experiment, as it was shown that the subject could discriminate between two realistic stimuli. The graph (Figure 8.1) shows that the distinction between $S\Delta$ and S^d in the present experiment was extremely strong.

However, once the treatments were removed, the strength of the learning was not clear. Although stereotypy was lower under $S\Delta$ in the return to baseline, it was not as suppressed as it had been during treatment, and it appeared that a gradual return to baseline levels would take place after some fluctuation in the behaviour. However, the discrimination which was achieved during intervention was strong enough to suggest that a more durable response could be achieved if generalisation techniques were incorporated into the programme.

The treatment package used in $S\Delta$ was highly effective in reducing the stereotypy, but the role of the individual components was not clear. The partial component analysis which was designed to assess the effect of momentary DRO was partly confounded by an unforeseen time of day effect. Despite this, the results of the follow-up strongly suggest that with high rate behaviour, momentary DRO might be facilitated by pre-trial instructions and (possibly) reprimands, at least initially, to suppress the behaviour and allow momentary DRO to operate. Whole-interval DRO has often been used as part of a treatment package, and this experiment shows that momentary DRO can be used in the same way. This finding may be extremely useful for carers who wish to combine an effective suppressive technique with other treatments.

Discrimination training, supported by a full treatment package as described in this present experiment, appears to be a useful technique for selectively reducing stereotyped

behaviour. However, a more thorough package might ensure a more durable effect. Firstly, the period of discrimination training would need to be extended, and there should be concurrent sessions in applied settings also. Secondly, the stereotyped behaviour would need to be replaced by a more appropriate behaviour which could perform the same function in a more acceptable fashion. The potential role of discrimination training in the social integration of people with such behaviours warrants further exploration.

Chapter 6

The relative accuracy and ease of use of momentary DRO and whole-interval DRO

This chapter describes two group experiments which were designed to simulate the way a care-worker unfamiliar with behavioural techniques might use a DRO schedule. These experiments explored one of the main aims of this thesis: to determine if momentary DRO is easier to use than whole-interval DRO.

Student subjects were taught to use a DRO schedule and give 'reinforcers' to an actor on a video screen whenever he fulfilled the requirements of the schedule. Half the subjects were required to programme whole-interval DRO while the other half programmed momentary DRO. In the first experiment, half of the subjects in each of the two DRO groups experienced distracting conditions. In the second experiment all subjects were required to complete a concurrent task. The purpose of these two experiments was to explore the relative ease of use and accuracy of whole-interval DRO and momentary DRO under different conditions.

The number of reinforcers awarded by subjects in each schedule was recorded, in order to provide further evidence for the different levels of reinforcement seen in momentary DRO and whole-interval DRO.

EXPERIMENT 9

The relative accuracy and ease of use of momentary DRO and whole-interval DRO under silent and distracting conditions

INTRODUCTION

One of the main aims of this thesis was to determine whether momentary DRO is easier to use than whole-interval DRO. It is hypothesised that it *is* easier, because it does not require continual concentration on the part of the user. The evidence for this comes from a number of sources, some of which were outlined in Chapter 1. For example, Tierney, McGuire, and Walton (1979) reported that care-staff had great difficulty implementing a whole-interval DRO schedule: "The degree of vigilance required by this schedule was considerable, and without exception each staff member reported difficulty in concentrating to the required degree over the 140 minutes" (p. 178).

Jones and Baker (1988a) and Saloviita (1988) encountered similar problems with the Differential Reinforcement of Incompatible behaviour (DRI) schedule, which like whole-interval DRO, requires continuous subject monitoring. In reviews of DRO studies, Poling and Ryan (1982) and Cooper (1987) highlighted the difficulty of implementing schedules which require a behaviour to be constantly observed. The same view was expressed more recently by Vollmer et al. (1993). It is also the experience of the author that momentary DRO allows simultaneous involvement in conversation and other tasks to a far greater extent than whole-interval DRO.

If momentary DRO requires less concentration than whole-interval DRO, then it may also be more accurate, because if attention wanders it is less likely to result in errors. Additionally, momentary DRO may be more resilient than whole-interval DRO under

conditions of noise and disruption. This is because it should be possible to carry out other tasks and permit interruptions whilst using momentary DRO in a way that is not possible in whole-interval DRO.

Most studies of DRO take place in very controlled, quiet conditions, and so the extent to which whole-interval DRO is effected by disruption has probably not been clearly seen. As whole-interval DRO is advocated so widely, it is important to find out if staff can implement it easily and accurately under the type of conditions in which it is actually used. If it is found that momentary DRO is easier and/or more accurate, this would have implications for clinicians' choice of treatment programmes.

This experiment, therefore was designed to test the following hypotheses:

1.   That subjects find momentary DRO easier to use than whole-interval DRO.

2.   That momentary DRO is implemented more accurately than whole-interval DRO.

3.   That momentary DRO is more resilient than whole-interval DRO under distracting conditions.

4.   That subjects using momentary DRO will award significantly more reinforcers than subjects using whole-interval DRO.

METHOD

Subjects

The subjects in this experiment were 60 volunteers recruited from a student subject pool and via a request sent on electronic mail. There were 17 males and 43 females and their ages ranged from 18 to 48. All subjects were undergraduate or postgraduate psychology students at University of Wales, Bangor. Undergraduates earned course credits, and postgraduates earned £3.00, for their participation. None of the subjects were familiar with the DRO schedules prior to the experiment.

Apparatus and task

The main apparatus in this experiment consisted of a 20-inch screen television, connected to a video-recorder, and an Apple Macintosh SE computer. The computer screen displayed a stop-watch which had been programmed in Hypercard for this experiment. The stop-watch was operated by three keys on the computer keyboard: 'S' for stop and start, 'R' for reset, and the space-bar for 'reward'. The three keys on the key-board had their stop-watch functions clearly written on them. The stop-watch program presented a screen showing large numbers (00.00, when first opened) and a drawing of a stop-watch with a reminder of the keys for operating it. When the 'S' key was pressed, the stop-watch numbers would start to move forward, one second at a time, as per a conventional hand-held stop-watch. If the 'S' key was pressed a second time, the numbers would stop, and remain at the last number until the 'R' (reset) key was pressed, when they would go back to zero (00.00). When the space-bar was pressed, a box containing the word 'Reward' would appear briefly at the bottom right-hand corner of the screen. The program could be set up to run a 'practice' or a 'trial' session. The only difference between the sessions was that nothing was recorded during the practices, while during trials, every key-press of the subject - and the time each key was pressed - was recorded.

The video played a tape of an actor modelling stereotyped behaviour. It was felt to be more appropriate to ask an actor to model stereotypy than use a film of a client, because the client would be unable to give their consent to be filmed. The actor, who was a care-worker at a local learning disabilities unit, was asked to model a stereotyped behaviour of a client who rubbed saliva into his hand. This behaviour was modified for the actor, who was asked just to model the hand-rubbing part of the stereotypy, as it was felt unacceptable to ask him to rub saliva into his hand. This stereotypy was chosen as appropriate for this experiment because of its distinctiveness, and for the ease with which subjects could determine its occurrence. Prior to the experiment, thirty minutes of recorded observation of the real client took place, and the stopping and starting times of his stereotypy were programmed into a computer. Thus when the actor was filmed, his demonstration of the behaviour was

guided by prompts on a computer which told him when to stop and start the stereotypy. In this way the modelling was as similar to the real behaviour of the client as possible. The actor, when not hand-rubbing, would engage in frequent head-swaying and vocalising, so that it did not appear as if he was simply waiting for his next hand-rubbing cue.

The actor was presented to subjects as 'Danny'. They were told that he was an actor. The film was edited into three sections: a 35 second segment designed to show 'Danny' to the subjects and show them what the hand-rubbing behaviour looked like; a five minute segment which was to provide subjects with a practice session; and a fifteen minute segment which was the main experimental trial session. Before both the practice and the trial segments, a prompt appeared ('Practice' or 'Start of Experiment'). Then at one second intervals the following sequence would appear: 'Ready', '5', '4', '3', '2', '1'. After '1' 'Danny' would appear on screen. The practice session ended after exactly five minutes with the words 'End of Practice'; the experiment trial ended after exactly fifteen minutes with the words, 'End of Experiment'. A clock ran at the bottom of the screen during both practice and experiment sessions so that subjects could see how much longer there was left for the experiment to run.

Prior to the experiment, the inter-response times of 'Danny's' hand-rubbing stereotypy were calculated (both the practice and experiment segments of film were used in this calculation). The average length of time between displays of the stereotypy was 44.2 seconds, and as it is customary to set a DRO interval of slightly less than the average IRT, a DRO interval of 35 seconds was decided upon.

Other materials used in this experiment were pre-experimental information sheets which were designed to give subjects background details of the experiment, and two experimental instruction sheets, one with full instructions and one with a summary of the instructions. There were two sets of these, one for momentary DRO and one for whole-interval DRO (see Appendix D). Subjects after the experiment were asked to fill in a simple

questionnaire (see Appendix D) which asked them to indicate how difficult they found it to carry out the schedule. The options were 'very easy', 'easy', 'quite easy', 'quite difficult', 'difficult', 'very difficult'. The questionnaire also asked subjects to give more detailed information about their thoughts on what they were asked to do in this experiment. A space was provided for these comments.

In order to simulate a real-world setting, where conditions can be noisy and distracting, a series of distractions was implemented for half the subjects. The distractions were scripted so that they were identical for all subjects. There were four distractions - two incoming telephone calls and two occasions where a confederate came into the room and interacted with the experimenter. The number and type of distractions was based on an audio-tape recording of a standard day-room in a local hospital for people with learning disabilities (see Appendix D for scripts).

Design

This experiment used a simulated situation in which a subject took the role of a care-worker using a DRO schedule. The experiment employed a two-way independent-subjects design, where subjects were randomly allocated to one of four conditions:

1   Whole-interval DRO   - silent room

2   Momentary DRO   - silent room

3   Whole - interval DRO - with distractions

4   Momentary DRO   - with distractions

Fifteen subjects were in each of these four conditions. Subjects were not aware that there were different conditions.

There were two independent variables:

(i)   The type of schedule (whole-interval DRO or momentary DRO)

(ii)   The condition (silent or distracting).

The three dependent variables were:

(i)   Subject's reported difficulty of task, as measured by a six-point Likert scale ranging from 'very difficult' to 'very easy'

and two measures of accuracy:

(ii)   The number of times subjects pressed the 'reward' key during the experiment

(iii)   The times at which subjects pressed the 'reward' key.

Procedure

Subjects were seated at a desk containing the computer, television and video. The television was to the left of the subject and the computer was placed in front of them. The video was on the right hand side of the desk. Subjects were given the experimental instructions to read, according to whether they were in the whole-interval DRO or momentary DRO condition. They were then shown a brief clip of 'Danny' so that they could gain an understanding of what he and the stereotyped behaviour looked like. Subjects were then given a second sheet, a summary of the experimental instructions.

In the whole-interval DRO condition, subjects were instructed to watch 'Danny' on the video, time him with the stop-watch whenever he ceased his stereotyped behaviour, and press 'reward' if he refrained from the stereotypy for 35 seconds. Subjects were instructed that if 'Danny' started hand-rubbing while they were timing him, they should stop the stop-watch, reset it to zero, and wait for him to cease before starting timing again. As a further reminder, they were told that they should never be timing 'Danny' when he was rubbing his hand.

In the momentary DRO condition subjects were told to start the stop-watch as soon as 'Danny' appeared on the screen, and to time him regardless of his stereotyped behaviour. Every 35 seconds they should stop the watch and check him. If he was displaying the stereotyped behaviour at that moment they should reset the stop-watch to zero and start

timing another 35 seconds. If he was not displaying the stereotyped behaviour at that moment, they should press 'reward', then reset the stop-watch and start timing again.

Once subjects had read the instruction sheets, the experimenter then explained how the stop-watch operated. The experimenter gave the subjects practice at using the keys by asking them to start the watch, let it run for ten seconds and then reset it to zero. Once subjects had done this, the experimenter asked them to start the watch again, run it again for ten seconds, and then press 'reward' and reset the watch to zero.

Subjects were then given a five minute practice. They were told that they could have the summary instruction sheet in front of them during the practice, though not during the experiment. It was explained that the experimenter would sit by them during the practice to make sure they were performing the task correctly. All subjects were told to try and respond as accurately and quickly as possible, and that the computer would not record anything they did during the practice. It was explained that the video would do a five second count-down and then 'Danny' would appear. During the practice, the experimenter initially prompted the subject to press the keys in the right order.

Once the subject had indicated that they were ready to begin the experimental trial, the experimenter explained that the video would once again count down five seconds before 'Danny' appeared. During the trial, the experimenter sat at a desk four feet away from the subject.

During the silent conditions the room was silent for the fifteen minutes of the trial. However, during the distracting conditions, the experimenter redirected the telephone to the confederate, who was situated in the room next door, phoning her first with a standard remark which was the signal for the confederate to start timing (see script in Appendix D).

Three minutes after the confederate received the notifying call, she made the first telephone call. One minute after this call she knocked on the door of the experimental room, entered, and she and the experimenter held a conversation, both standing behind the subject. After this conversation ended the confederate returned to the room next door. Three minutes later she made the second telephone call. Three minutes after this call she again knocked on the door of the experimental room, where a brief exchange with the experimenter took place.

At the end of the experiment all subjects were asked to fill in a copy of the questionnaire.

After the experiment was completed, debriefing sheets explaining its purpose were sent out to all subjects.

RESULTS

Difficulty of task

A two-way analysis of variance was carried out on subjects' difficulty of task scores. There were two between subjects factors - DRO schedule and condition. The means can be seen in the tables below.

Table 9.1 Means and standard deviations of subjects' recorded difficulty of task across whole-interval DRO and momentary DRO

	Whole-interval DRO	Momentary DRO
Mean	3.93	3.90
Standard deviation	0.94	0.92

Table 9.2 Means and standard deviations of subjects' recorded difficulty of task across whole-interval DRO and momentary DRO and across silent and distracting conditions

	Whole-interval DRO		Momentary DRO	
	Silent	Distracting	Silent	Distracting
Mean	3.87	4.00	3.87	3.93
Standard deviation	1.06	0.85	0.99	0.88

The ANOVA showed that there was no difference in task difficulty between whole-interval DRO and momentary DRO. There was also no effect of the silent and distracting conditions. The means for both whole-interval DRO and momentary DRO were around 3.9, suggesting that subjects found both schedules easy (4 on the Likert scale was equivalent to a response of 'easy').

Qualitative responses concerning the experiment

Subjects were asked to give some more detailed information about their thoughts on what they were asked to do in the experiment. All subjects commented. Counts were made of the number of subjects who said that the experiment was 'easy', or 'easier' than they thought, or 'straightforward', and counts were also made of the number of subjects who said that the experiment was 'hard' or 'difficult'. The results can be seen in the table below.

Table 9.3 Number of subjects who said they found the experiment easy or difficult in their qualitative responses

	Whole-interval DRO		Momentary DRO	
	Silent	Distracting	Silent	Distracting
Subjects who said 'Easy', 'Easier', 'Straightforward'	4	6	4	9
Subjects who said 'Hard', 'Difficult'	5	2	3	2

Chi-square tests across whole-interval DRO and momentary DRO, across whole-interval DRO silent and whole-interval DRO distracting, and across momentary DRO silent and momentary DRO distracting all showed non-significant results. In other words, there was no difference between or within schedules with regards to subjects' describing the task as easy or difficult.

The results for momentary DRO silent and whole-interval DRO silent were combined into one group, and the results for momentary DRO distracting and whole-interval DRO distracting into another group. A Chi-square test showed that there was a significant difference between groups at the 10% level ($X^2 (1, N=35) = 3.23, p<0.1$). This suggests a trend that, regardless of schedule, there was a difference in reports of 'easy' and 'difficult' across silent and distracting conditions. Closer analysis of the data reveals the unexpected result that there were almost twice as many mentions of 'easy' by subjects in distracting conditions than by subjects in silent conditions (15 as against 8 respectively). Conversely, there were exactly half the number of mentions of 'difficult' by subjects in the distracting conditions than by those in the silent conditions (4 against 8).

Most of the other comments were made by just one or two subjects within each condition. The only exception to this was that five subjects in the momentary DRO silent condition reported finding concentration difficult or feeling their mind wandering. Similar comments were made by four subjects in the whole-interval DRO silent condition, two subjects in the whole-interval DRO distracting condition and one subject in the momentary DRO distracting condition.

Accuracy of the number of reinforcers
The number of reinforcers subjects gave was analysed using a two-way ANOVA. There were two between subjects factors: DRO schedule and condition. A highly significant difference was found between schedules ($F(1, 56) = 960.378, p<0.001$). A mean of 12

reinforcers was given in whole-interval DRO, and a mean of 16.4 in momentary DRO. This difference can be clearly seen in Figure 9.1.
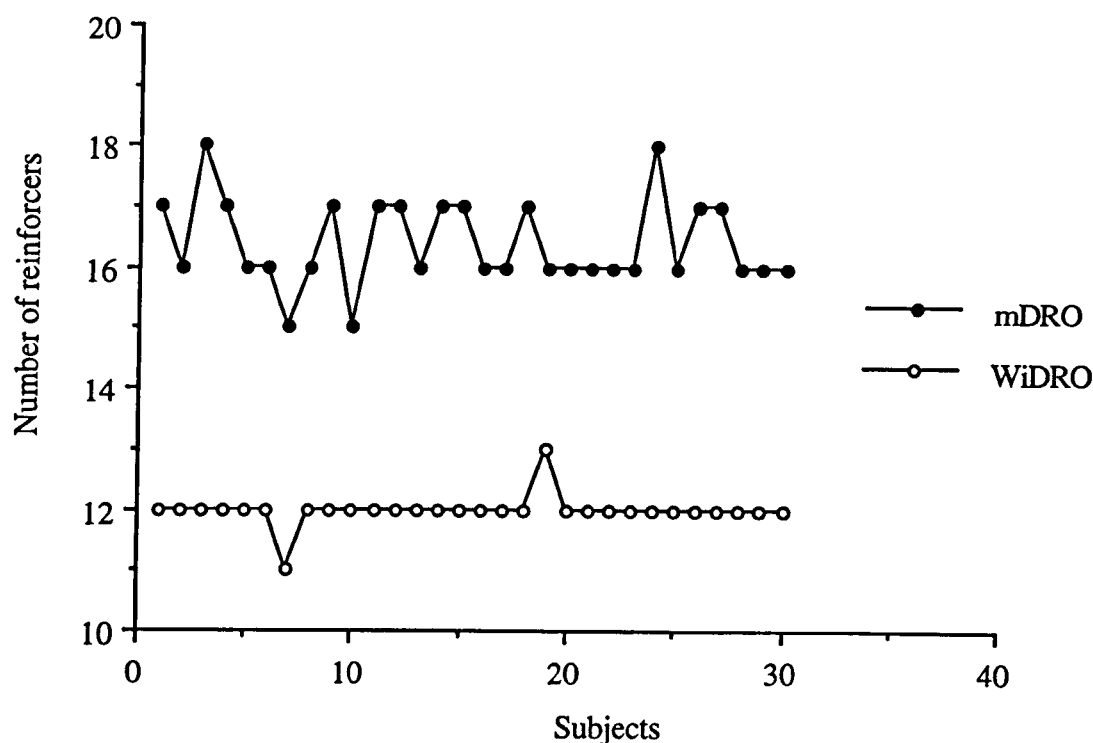
Figure 9.1 : Number of reinforcers given in momentary DRO and whole-interval DRO.

Because of the large discrepancy between the number of reinforcers given in the two schedules, it was necessary to calculate a measure that showed how far subjects deviated from the correct number of reinforcers, whilst taking into account the fact that the two schedules were operating with completely different rates of reinforcement. This was done by dividing the number of reinforcers subjects gave by the number that they should have given (calculated to be 12 in whole-interval DRO and 16 in momentary DRO). This established a percentage error score for each subject. These percentage error scores were analysed using a two-way analysis of variance. The two between subject factors were DRO schedule and condition. The ANOVA showed that there were significantly more errors in momentary DRO than whole-interval DRO ($F_{(1, 56)} = 12.218$, $p < 0.001$). In

other words, considerably more subjects gave the correct number of reinforcers in whole-interval DRO than in momentary DRO. This can be seen in Figure 9.1, where there is much less variability in whole-interval DRO than in momentary DRO.

Overall, six times more errors were made in momentary DRO than in whole-interval DRO (the means were 3.33% for momentary DRO and 0.56% for whole-interval DRO). Figure 9.2 clearly illustrates the greater errors made in momentary DRO.
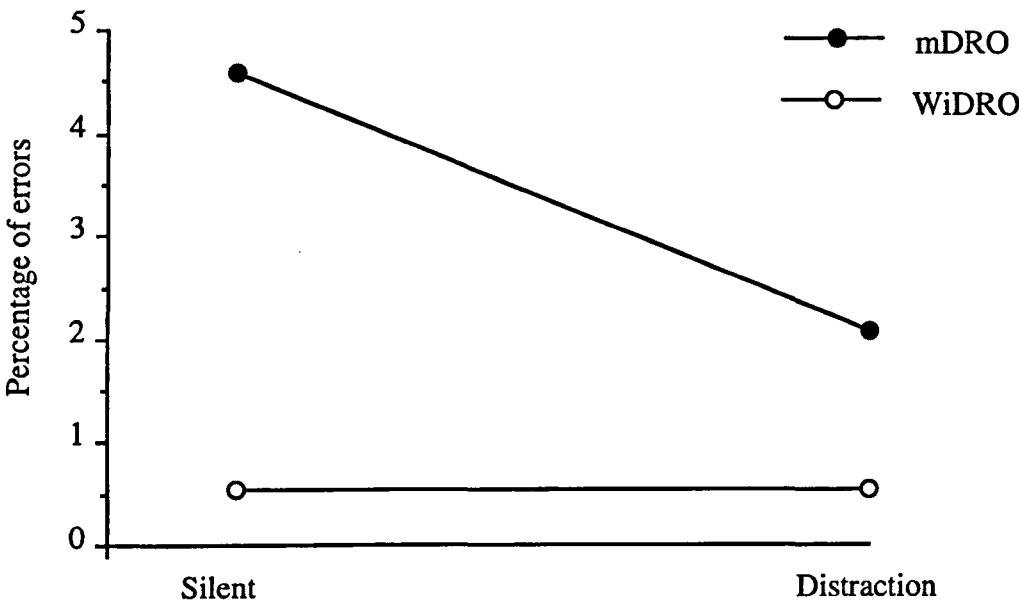
Figure 9.2 : Percentage of errors made on number of reinforcers across schedule and condition

The larger number of errors in momentary DRO can also be illustrated by comparing the standard deviations of the actual number of reinforcers given in momentary DRO and whole-interval DRO (0.72 and 0.26 respectively). There is a far higher variability in momentary DRO with regard to the number of reinforcers given.

The effect of distractions on the number of reinforcers

There was no effect of the silent or distraction condition on accuracy in whole-interval DRO (the mean percentage errors were identical in silent and distracting conditions). However, in momentary DRO there were over twice as many mistakes in the silent condition than in the distracting condition. A t-test showed that this difference did not quite reach significance at the 5% level ($t(28) = 1.809$, $p<0.1$, two-tailed). Figure 9.2, above, illustrates this trend.

Accuracy of the timings of reinforcers

In addition to the *number* of reinforcers which subjects gave, a second measurement of accuracy was calculated, based on the *times* at which the reinforcers were given. A standard set of times at which reinforcers should be given in each of the two schedules was determined, and from this an 'acceptable range' for each reinforcer was calculated. An 'acceptable range' was defined as the subject giving the reinforcer within two seconds before the optimal time, or up to four seconds after the optimal time. Such responses were classed as correct. Any response outside this range was classed as incorrect. The two seconds on either side of the correct time were intended to allow for any slight error in timing from the start of the experiment. The additional two seconds allowed after was so that responses could be classed as correct if they were given just a little late, but near enough to the correct time that the contingency and reinforcer were in close proximity.

The tables below show the number of subjects in each condition who gave all the reinforcers within an acceptable time range; the number of subjects who delivered between 50% and 99.99% of the reinforcers within an acceptable time range; the number of subjects who delivered between 1% and 49.99% of the reinforcers within an acceptable time range; and the number of subjects who failed to give any of the reinforcers within an acceptable time range.

Table 9.4 Numbers of subjects who gave reinforcers within acceptable time ranges across schedule

	Whole-interval DRO	Momentary DRO
No. Ss who were 100% correct	19	8
No. of Ss between 50% and 99.99%	11	9
No. of Ss between 1% and 49.99%	0	10
No. Ss who were 0% correct	0	3
TOTAL SUBJECTS	*30*	*30*

Table 9.5 Numbers of subjects who gave reinforcers within acceptable time ranges across schedule and condition

	Whole-interval DRO		Momentary DRO	
	Silent	Distracting	Silent	Distracting
No. Ss who were 100% correct	10	9	4	4
No. of Ss between 50% and 99.99%	5	6	3	6
No. of Ss between 1% and 49.99%	0	0	7	3
No. Ss who were 0% correct	0	0	1	2
TOTAL SUBJECTS	*15*	*15*	*15*	*15*

It can be seen from these tables that considerably more subjects in the whole-interval DRO conditions gave reinforcers within an acceptable time range than subjects in the momentary DRO conditions. A Chi-square test was carried out on the number of subjects who achieved above 50% accuracy and those who achieved below 50% accuracy across both whole-interval DRO and momentary DRO (regardless of silent or distracting conditions). This showed a highly significant difference between the two schedules, with many more subjects achieving above 50% accuracy in whole-interval DRO than in momentary DRO $(X^2 (1, N=60) = 16.596, p<0.001)$. Indeed, all the subjects in whole-interval DRO gave

upwards of 50% of reinforcers within an acceptable time range, compared to just two-thirds of subjects in momentary DRO.

For each subject, a measure of overall accuracy of timing was calculated by determining the percentage of reinforcers which were given within an acceptable time limit. The table below shows the means and standard deviations of these percentage scores.

Table 9.6 Means and standard deviations of percentage accuracy of timing scores across whole-interval DRO and momentary DRO

	Whole-interval DRO		Momentary DRO	
	Silent	Distracting	Silent	Distracting
Mean	93.33%	92.78%	54.58%	61.67%
Standard deviation	10.06	10.85	41.76	42.25

The percentage scores were analysed using a two-way analysis of variance. The two between subjects factors were DRO schedule and condition. The ANOVA showed that the difference between the two schedules was highly significant ($F(1, 56) = 19.535$, $p<0.001$). That is, subjects timed a far higher percentage of reinforcers correctly in whole-interval DRO than in momentary DRO.

Effect of distractions on timing of reinforcers

There was no effect of the silent or distracting conditions in whole-interval DRO. There was a small difference between subjects in momentary DRO, as can be seen in Table 9.6: accuracy in timing of reinforcers was slightly greater in the momentary DRO distracting condition than the momentary DRO silent condition. However, a t-test showed that this difference was not significant.

These results confirm the findings of accuracy as measured by number of reinforcers: that subjects in whole-interval DRO were more accurate than subjects in momentary DRO, and that trends suggest that subjects in the momentary DRO silent condition were less accurate than subjects in the momentary DRO distracting condition.

Correlations between the two measures of accuracy

Two different measures of accuracy were used in this experiment: the percentage error scores for number of reinforcers and the percentage of accuracy in timing the reinforcers. Two correlations were carried out to examine the nature of any relationship between these two measures. The first correlation was carried out between percentage error scores and percentage accuracy of timing for whole-interval DRO. A significant negative correlation was shown $(r(28) = -0.587, p<0.01)$. This means that in whole-interval DRO, as the accuracy of timing reinforcers went down, so the percentage of errors in the number of reinforcers went up. The second correlation was carried out between percentage error scores and percentage accuracy of timing for momentary DRO. No relationship was found between the two measures.

Therefore there was a relationship between the two measures of accuracy in whole-interval DRO but there was no such relationship in momentary DRO.

DISCUSSION

Three of the hypotheses proposed at the start of this experiment were not supported. Subjects did not report that momentary DRO was easier to use than whole-interval DRO. Furthermore, subjects in momentary DRO were significantly less accurate than subjects in whole-interval DRO. There were no statistically significant effects of distractions, although there was a small trend which suggested that subjects in momentary DRO were more accurate under distracting conditions than silent conditions. Support was found for the

fourth hypothesis - subjects using momentary DRO did award significantly more reinforcers than subjects using whole-interval DRO.

Difficulty of task

Subjects found both schedules equally easy to use. It is possible that this was because the situation in this experiment was not analogous with that of a care-worker. Although some of the subjects were exposed to distractions, they were not obliged to engage in them. They did not have to carry out any other tasks that might take them away from the main task. In an attempt to simulate a higher demand situation a further experiment was carried out, where subjects were required to carry out a concurrent task while running the DRO schedule (see Experiment 10 later in this chapter).

The trend in the qualitative responses which suggests that overall, subjects in distracting conditions perceived their task as easier than subjects in the silent conditions is in line with some of the other findings (whereby, for example, accuracy improved in momentary DRO under distracting conditions). There are two possible explanations for this trend. The first is that the type of distractions used in this experiment concentrated the subjects' minds. Subjects in all conditions reported that they found themselves losing concentration; however there were more of these reports in the silent conditions than in the distracting conditions. It may be that in the distracting conditions the torpor induced by the task (and by the hypnotic vocalising of the actor on the video), was enlivened by the various interruptions and helped to keep the subjects alert. It may be that rather than distracting the subjects from their task, the interruptions actually achieved the opposite result. This point is also discussed below in relation to the accuracy of reinforcer findings.

The second explanation concerns the social effect of the distracting conditions. Subjects may have been discomfited or even embarrassed by the distractions and so tried harder to do the task well and spoke more positively of it afterwards. Very few subjects mentioned the distractions when they were asked to fill in the questionnaire after the experiment, or

when they were given the opportunity to ask questions. There is a small amount of evidence to suggest that some subjects did regard the distractions as unprofessional interruptions to the experiment. Two or three subjects mentioned them spontaneously some time after the experiment, and one subject asked another subject whether she thought that the confederate who had been doing the interrupting was perhaps 'a bit strange' to persistently ignore the 'do not disturb' sign on the door of the experimenting room.

There was little response to the debriefing sheets which were sent to all subjects after the experiment had been completed. These asked subjects who had been in distracting conditions if they had been aware of the disruptions. One subject did send a written response, 'Yes, I did notice the distractions, and yes, they did bother me slightly. However, I don't think they made that much difference to my performance, but then again I was concentrating much harder in that situation than I probably would be in a real life situation.'

Accuracy of the number and timing of reinforcers

In line with previous studies in this thesis, more reinforcers were delivered in the momentary DRO conditions than in whole-interval DRO. It was possible for 'Danny' (the actor on the video) to earn twelve reinforcers in whole-interval DRO, and sixteen in momentary DRO.

Considerably more subjects gave the correct number of reinforcers in whole-interval DRO (93%) than in momentary DRO (33%). Part of the explanation for this is that momentary DRO relies on a 'snap-shot' of the target behaviour. Most of the time, the client is not observed; it is only at a particular moment that a decision about reinforcement is made. It is possible to see how, if a subject started timing at slightly the wrong moment - perhaps just two or three seconds out - how this might impact on later decisions about reinforcement. For example, if a subject started timing at the correct moment, then 35 seconds later 'Danny' might not be displaying the stereotyped behaviour, and so he would receive a

reinforcer. However, if they started timing two seconds too soon or too late, they would have to make a reinforcement decision on what was effectively the 33rd or the 37th second, and at these moments 'Danny' may be displaying the stereotypy. In contrast, whole-interval DRO requires continual observation and a reinforcer is not awarded unless the behaviour has been absent for the whole 35 seconds. Therefore a few seconds too early or too late has less effect on the reinforcement decision, as the decision is not made on the basis of just a moment but on much more information.

Another effect of using such a small amount of information to make a decision in momentary DRO is that it is simply easier to make a mistake. If the subject had not been thinking about the task during an interval, then as their attention was redirected to the task there were several ways in which mistakes could occur. They might get flustered and inadvertently press the wrong keys, or they might temporarily forget whether they were meant to reinforce the absence or presence of the stereotypy. Or, if the client was not hand-rubbing at the 35th moment, but began shortly afterwards, a subject who had hesitated at the crucial moment might not like to press the 'reinforcer' key as the behaviour had by now re-started. This could also happen the other way round: the client might have been hand-rubbing at the 35th moment but stopped shortly afterwards. A subject who hesitated might then press the 'reinforcer' key inappropriately. Alternatively, a subject might fail to turn to the stop-watch at the correct moment, and then once the 35th second had passed, the subject would not be sure whether or not they should have given a reinforcer.

Subjects in momentary DRO mentioned all these ways of making mistakes after the experiment. It would be much harder to make similar mistakes in whole-interval DRO if attention did not wander, as subjects were required to watch the client continuously and thus had far greater amounts of information on which to base decisions about reinforcement.

Consideration must also be given to the cumulative effect of mistakes in momentary DRO which is caused by the way the intervals in this schedule is timed. The timing in momentary DRO is continuous, apart from the brief stopping and resetting every 35 seconds; therefore, a wrongly timed reinforcer will necessarily effect the timings of all subsequent reinforcers. If a subject stops timing an interval on the 37th second instead of the 35th, then the next interval will start at the wrong time, and so will the one after that, and so on. Even very small errors in timing will impinge on all later timings. However, in whole-interval DRO timing is not a continuous process but is composed of many stops and starts determined by the target behaviour. In this sense, the timing and reinforcing in whole-interval DRO can be seen as discrete units. A wrongly timed reinforcer, therefore, need not effect any subsequent timings because although a subject might restart the stop-watch at the wrong moment, if the target behaviour occurs before the end of the interval then the subject stops and resets. When the behaviour ceases and the subject restarts the watch, they are at the same point as other subjects who had given the previous reinforcer correctly. If there is no occurrence of the target behaviour in the interval after a wrongly timed reinforcer, then it is possible that a subsequent reinforcer will be wrongly timed. But any opportunity that arises to stop and reset the watch will allow the timings to come back into line.

The difference in the timing of intervals may explain the pattern of errors shown by subjects in this experiment. In momentary DRO, the point where an error in timing was made subsequently effected all other timings so that they were all wrong. Twenty-two subjects in momentary DRO made an error in timing at some point. Of these, fifteen subjects remained wrong for the rest of the timings. Therefore, only seven subjects managed to make good their errors and give some later reinforcers accurately. Of these seven, only two made a complete recovery (e.g. made just one mistake and then performed correctly for the remainder of the experiment). In contrast, eleven subjects in whole-interval DRO made an error in timing at some point. Of these, nine managed to recover their performance totally, so that after some early mistakes, the remainder of their

reinforcers were timed correctly. The remaining two subjects each made errors in two separate places.

Therefore it appears that in momentary DRO, most subjects time a reinforcer incorrectly at some point and that all subsequent decisions are then taken at the wrong time, thus effecting the number of reinforcers given and the times at which they are given. In whole-interval DRO, on the other hand, some subjects make mistakes but the nature of the schedule allows them to make good their error at a later stage.

To sum up: because one relies on very little information in momentary DRO on which to base a reinforcement decision - just a moment's observation - it is easier to make a mistake in the first place than in whole-interval DRO. Then, the fact that subsequent timings are effected by earlier ones means that it is easier to continue making mistakes in momentary DRO.

Correlations between measures of accuracy

Two different measures of accuracy were used in this experiment. The first was the percentage error of the number of reinforcers. The second was the accuracy of timing of reinforcers. It was necessary to find out whether these two measures were providing essentially the same information. If the measures were closely related this would provide some useful information about the workings of the schedules. A strong negative relationship between measures might be expected - the less accurate the timing of reinforcers, so the greater the error on number of reinforcers.--This relationship was shown between the measures for whole-interval DRO. However, no such relationship was shown for momentary DRO. This provides more information to suggest that the two schedules operate in different ways.

Presumably the reason why there was a relationship between timing and errors in whole-interval DRO and not in momentary DRO was because of the difference outlined above

regarding the timing of intervals. Once the timings are wrong in whole-interval DRO, more errors are made regarding the number of reinforcers, but there is not a similar relationship between number and timing of reinforcers in momentary DRO. This shows that although mistakes in timing in momentary DRO have a subsequent effect on all other timings, mistakes do not make it more likely that subjects will deliver an incorrect number of reinforcers. Conversely, although mistakes in timing do not necessarily impact upon later timings in whole-interval DRO, they do impact on accuracy of reinforcers - presumably the impact is on the reinforcer that is wrongly given or withheld because of that particular wrong timing rather than any subsequent ones. That is, if one makes a mistake in timing in whole-interval DRO one is quite likely to make a concurrent mistake in giving a reinforcer, *because of* the error in timing.

Effect of distractions on accuracy

The accuracy of subjects in distracting conditions was not impaired in either schedule. In whole-interval DRO subjects performed equally in both conditions. In momentary DRO there was a trend which suggested that subjects performed slightly more accurately under distracting conditions than under silent conditions. This trend was also reflected in the qualitative responses regarding difficulty of task.

Some possible reasons for this trend have already been outlined above in relation to subjects' reported difficulty of task, and can be applied to the accuracy outcomes. It may be that this trend was not seen in whole-interval DRO because so few errors of any kind were made - indeed, so few that there was almost a ceiling effect of high performance. Thus, any positive effect of the distractions in whole-interval DRO were not seen because performance could not be improved.

As mentioned previously, it is possible that subjects were made attentive by the interruptions rather than distracted by them. It may be that subjects were less bored during whole-interval DRO than in momentary DRO, and so did not need the enlivening effects of

the distractions. In whole-interval DRO subjects were 'on-task' all the time; they had to watch the video and keep attentive to the behaviour of 'Danny' and be ready to work the stop-watch. Subjects in momentary DRO, on the other hand, were not required to be vigilant and could spend periods of time 'off-task'. Once the stop-watch had been activated subjects had nothing to do for the next 35 seconds until the moment came to make a reinforcement decision and restart the stop-watch. It is possible that the nature of the student subjects (highly willing, eager to be good subjects) contributed to the results seen in this experiment. It had been anticipated that whole-interval DRO, requiring so much vigilance and concentration, would be perceived as difficult and indeed, would actually be more difficult to do well, than momentary DRO. However, subjects who were keen to perform well on the task may have responded positively to the high demands of whole-interval DRO, while subjects who were asked to carry out the momentary DRO schedule may have felt that they were not being required to do very much by only having to look at a television screen twice a minute. In short, the concentration required in whole-interval DRO may have come easily to these subjects while the more relaxed demands of momentary DRO might have led subjects to lose concentration. Five subjects in the momentary DRO silent condition mentioned losing concentration, compared to just one in the momentary DRO distracting condition, suggesting that the distractions aided concentration. Furthermore, nine subjects reported finding momentary DRO distraction 'easy' and just two reported finding it 'difficult', compared to more or less equal numbers who found momentary DRO silent 'easy' or 'difficult'. It is possible, therefore, that the distractions broke up the monotony in momentary DRO and kept subjects more attentive to the task.

Conclusion

Three of the hypotheses made in this experiment were not supported. Whole-interval DRO was found to be as easy as momentary DRO, and was also more accurate. Momentary DRO was not more resilient than whole-interval DRO under distracting conditions. The

fourth hypothesis - that more reinforcers would be delivered in momentary DRO than whole-interval DRO - was supported.

A number of reasons for these findings were considered. It is possible that the experimental situation was not similar enough to the environment it was aiming to simulate. The distractions in this experiment were closely modelled on some of the types of interruptions that can occur in care environments. However, there was one key difference: the distractions in the present experiment did not require the subject to engage in anything other than the task. The subject was not even required to look away from the screen and the stop-watch. In a 'real world' situation a care-worker would often have to engage with interruptions, perhaps by speaking to someone or attending to another client. Therefore it was decided that a further experiment of this kind should be carried out, where subjects were required to become actively engaged in the distractions by carrying out an additional task. This experiment is described below.

EXPERIMENT 10

The relative accuracy and ease of use of whole-interval DRO and momentary DRO when combined with a concurrent task

INTRODUCTION

This experiment was designed to investigate a matter arising from the previous experiment: namely, whether requiring subjects to actively engage in another task will have a differential effect in momentary DRO and whole-interval DRO. Experiment 9 found that external distractions had no effect on the accuracy of implementation of whole-interval DRO, and a slight effect of improving accuracy in momentary DRO. Overall, subjects implemented whole-interval DRO with considerably more accuracy than momentary DRO. It was decided to compare the two schedules under more distracting conditions than before, where subjects were required to engage in another activity concurrently with running the schedule. It has been reported that whole-interval DRO is difficult to carry out in conjunction with other tasks. Vollmer et al. (1993), for example, pinpoint the resetting element of whole-interval DRO as being particularly cumbersome, and state that such a procedure "..can be difficult for a parent with other household duties or staff with other clients to assist" (p. 10).

This experiment, therefore was designed to test the following hypotheses, based on the findings of Experiment 9:

1.   That when subjects are required to carry out a concurrent task, they find momentary DRO easier to use than whole-interval DRO.

2.   That when subjects are required to carry out a concurrent task, momentary DRO is implemented as accurately as whole-interval DRO.

It was decided that all subjects would be required to carry out the concurrent task, to allow a direct comparison between schedules. It was not necessary to carry out silent control conditions as subjects' performance in this experiment could be compared with subjects' performance in the silent conditions in Experiment 9. To facilitate this, it was decided to run thirty subjects in this experiment - fifteen in each DRO schedule - to allow for easy comparison across the experiments.

In an attempt to provide further support for the theory that momentary DRO provides higher levels of reinforcement than whole-interval DRO, a third hypothesis was tested, as follows:

3.   That subjects using momentary DRO will award significantly more reinforcers than subjects using whole-interval DRO.

METHOD

Subjects

The subjects in this experiment were 30 volunteers recruited from a student subject pool and via a request sent on electronic mail. There were 9 males and 21 females and their ages ranged from 18 to 48. All subjects were undergraduate or postgraduate psychology students at University of Wales, Bangor. Undergraduates earned course credits, and postgraduates earned £3.00, for their participation. Eight of the subjects had taken part in Experiment 9, and they were all allocated to the other schedule (either momentary DRO or whole-interval DRO) than the one they had used before. None of the other subjects were familiar with the DRO schedules prior to the experiment.

Apparatus and tasks

The apparatus used were the same as in Experiment 9. Experimental instructions included an additional paragraph which outlined the nature of the concurrent task. There were also

some additional materials, including the items used for the concurrent task. This consisted of 19 Information questions taken from the Revised Weschler Adult Intelligent Scale (British supplement, 1986). Four of the WAIS-R questions were used during the practice session, and fifteen during the trial session. Each of the questions was typed onto a separate half-sheet of paper with the number of the question typed clearly at the top of the sheet (e.g. 'Question 1'), with a space below the question headed with 'Answer:'. The questions themselves were presented in bold type. The 19 sheets were presented in a lever-arch folder which was placed, open, to the right of the subject on the same desk as the other apparatus. To prevent the file being moved, it was tied to the desk with a cord of the type used to secure computers to desks. A pen was placed at the top of the folder.

An audio-tape recorder was placed to the left of the subject. The tape-recorder contained a tape of the experimenter prompting the subject to turn to the folder and answer a question. The prompts were in the form of 'Question 1', 'Question 2', etc. There were 19 prompts, one for each question. Four of these prompts were during the practice session. To allow the subject time to become familiar with the DRO schedule, the first prompt was not until two minutes after the start of the session. The remaining 15 prompts were in the trial session. Seven of the gaps between prompts were less than one minute (30, 25, 40, 45 or 50 seconds) and seven of the gaps were greater than one minute (70, 75 or 80 seconds). This meant that a question occurred on average every minute, but there was not a regular pattern which might alert the subject to prepare to answer a question.

Design

This experiment employed an independent-subjects design, where subjects were randomly allocated to one of two conditions: whole-interval DRO or momentary DRO. Both conditions required subjects to carry out a DRO schedule whilst performing a concurrent task.

The independent variable was the type of schedule (whole-interval DRO or momentary DRO). The three dependent variables were the same as the previous experiment: subjects' reported difficulty of task; and the two accuracy measures, the number of reinforcers given and the times at which they were given.

Procedure

The procedure was largely as described in Experiment 9, other than the scripted distractions which were not used in this present experiment. Subjects were given the same experimental instructions as before, with an additional paragraph as follows:

'Throughout the experiment, a tape recorder will give you prompts at various intervals to answer a general knowledge question. The questions are in a folder on your right. When you hear the prompt, you should turn immediately to the folder and write your answer in the space provided.'

Subjects were given additional verbal information that when they heard the prompt they should turn immediately to the folder, and not turn back to the stop-watch and video until they had written an answer. The exception to this was if the stop-watch at the time of the prompt was extremely close to 35 seconds (e.g. from 32 seconds upwards); in such an instance the subject was permitted to complete the schedule requirements before turning to the folder.

Subjects were told that they could not move the folder, but should move towards it, either by reaching across or by moving their chair to the folder (the chair was on wheels). This was to ensure that subjects did not try to put the folder on the desk in front of them in an attempt to answer the question while watching the video. The first question was covered up by a blank sheet of paper. Subjects were told that on hearing the first prompt they should turn over that first sheet, write an answer to the question on the sheet underneath, and then turn back to the video without turning over the page for the next question. Thus

each time there was a prompt the procedure was to turn to the folder, turn the previous page over, write an answer on the next sheet and then turn back to the video.

Subjects were informed of the purpose of the experiment: that having to turn away from the video and engage in another task was designed to see how accurately they could continue to do the main task. Therefore subjects were asked to write something down in answer to each question, even if they didn't know the correct answer, so that at each prompt they temporarily disengaged from the main task. Subjects were also instructed that they should try and do both tasks to the best of their ability.

Subjects were given a five-minute practice at using the schedule and answering the questions. After the practice had been run, the trial began, and the experimenter sat at a desk four feet away from the subject. There was no contact between subject and experimenter during the fifteen minutes trial session.

As in Experiment 9, subjects were asked to complete a brief questionnaire at the end of the experiment.

RESULTS

Difficulty of task

Table 10.1 shows that the mean difficulty of task scores for subjects in both schedules were identical at 3.2. This indicates that subjects found both schedules quite easy (3 on the Likert scale was equivalent to a response of 'quite easy'). It was not necessary to carry out any statistical analyses on these scores as there was obviously no difference in difficulty between whole-interval DRO and momentary DRO.

Table 10.1 Means and standard deviations of subjects' recorded difficulty of task in whole-interval DRO and momentary DRO

	Whole-interval DRO	Momentary DRO
Mean	3.2	3.2
Standard deviation	1.01	1.08

Qualitative responses

All subjects wrote something, many commenting on carrying out two tasks at the same time. Counts were made of the number of subjects who said that the experiment was 'easy', or 'easier' than they thought, or 'straightforward', and counts were also made of the number of subjects who said that the experiment was 'hard' or 'difficult'. The results can be seen in Table 10.2 below.

Table 10.2 Number of subjects who said they found the experiment easy or difficult in their qualitative responses

	Whole-interval DRO	Momentary DRO
Subjects who said 'Easy', 'Easier', 'Straightforward'	5	7
Subjects who said 'Hard' or 'Difficult'	3	6

A Chi-square test across whole-interval DRO and momentary DRO was not significant. Thus there was no difference between schedules with regards to subjects describing the task as easy or difficult.

Most of the other comments were made by just one or two subjects within each condition. Comments made by three or more subjects included five subjects in momentary DRO who said that if the questions required any thought it was easy to lose track of the task. Four subjects in whole-interval DRO made similar comments. Three subjects in whole-interval

DRO said that they didn't feel that answering the questions was very distracting, and three subjects in momentary DRO made suggestions for how the task could have been more disruptive (e.g. by the use of multiple choice questions, by having the questions more closely spaced, or by answering the questions verbally). Three subjects in whole-interval DRO said they found it hard to do both tasks at once, and three subjects in momentary DRO said that they found the task difficult though they didn't specify whether it was the individual tasks they found hard or carrying them out together.

Accuracy of the number of reinforcers

The number of reinforcers subjects gave was analysed using a t-test. As in Experiment 9, a highly significant difference was found between schedules ($t(28) = -16.881$, $p<0.001$, two-tailed). That is, considerably more reinforcers were given in momentary DRO (a mean of 16.33 reinforcers) than in whole-interval DRO (a mean of 11.93). This is illustrated in Figure 10.1.
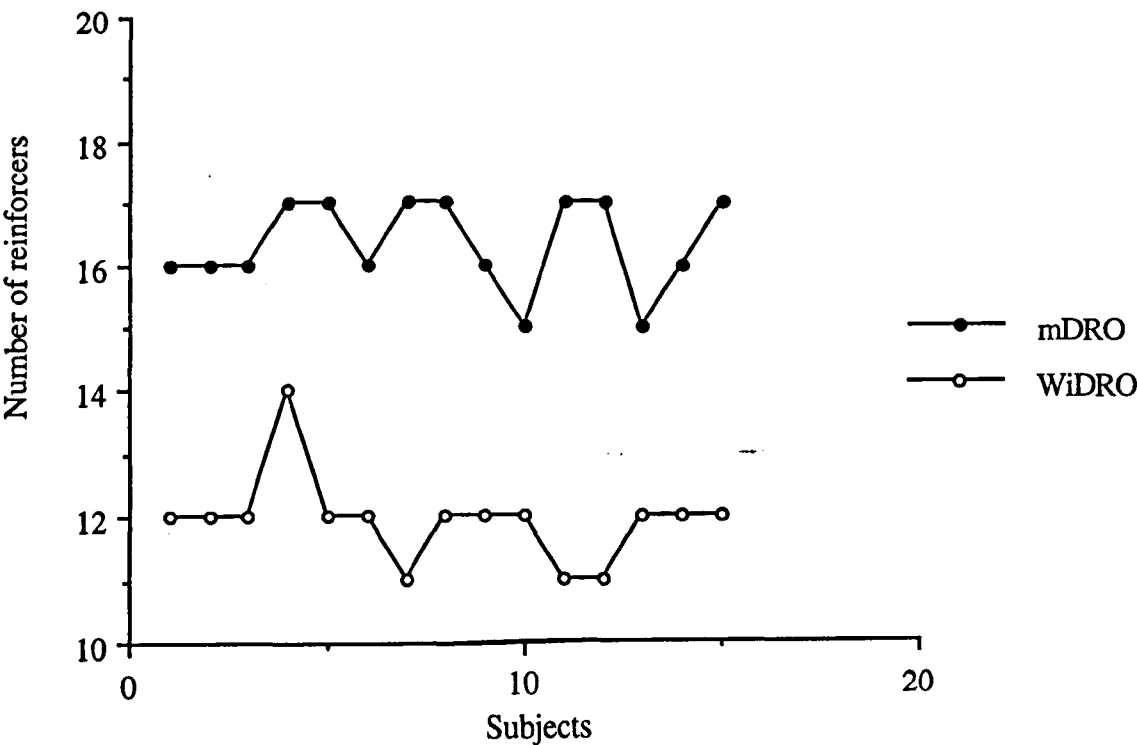
Figure 10.1: Number of reinforcers given in momentary DRO and whole-interval DRO

Percentage error scores were calculated for each subject, as in Experiment 9. There was a small difference in the overall mean number of errors made in momentary DRO (3.75%) and whole-interval DRO (2.78%). However, a t-test showed that this was not a significant difference. That is, an equivalent percentage of error was made in both schedules. Figure 10.1 depicts similar levels of variability in the two schedules.

Accuracy of the timings of reinforcers

Table 10.3 shows the number of subjects in each condition who gave all reinforcers within an acceptable time range; the number of subjects who delivered between 50% and 99.99% of the reinforcers within an acceptable time range; the number of subjects who delivered between 1% and 49.99% of the reinforcers within an acceptable time range; and the number of subjects who failed to give any of the reinforcers within an acceptable time range.

Table 10.3 Numbers of subjects who gave reinforcers within acceptable time ranges across schedule

	Whole-interval DRO	Momentary DRO
No. Ss who were 100% correct	4	2
No. of Ss between 50% and 99.99%	11	1
No. of Ss between 1% and 49.99%	0	10
No. Ss who were 0% correct	0	2

A Chi-square test was carried out on the number of subjects who achieved above 50% accuracy and those who achieved below 50% accuracy. This showed a highly significant difference between the two schedules, with many more subjects achieving above 50% accuracy in whole-interval DRO than in momentary DRO (X^2 (1, N=30) = 21, p<0.001). All the subjects in whole-interval DRO gave upwards of 50% of reinforcers within an acceptable time range, compared to only one-fifth of subjects in momentary DRO.

As in Experiment 9, percentage of accuracy of timing scores were calculated for each

subject. The means, standard deviations and ranges of these scores can be seen in Table

10.4.

Table 10.4 Means, standard deviations and ranges of percentage accuracy of timing scores across whole-interval DRO and momentary DRO

	Whole-interval DRO	Momentary DRO
Mean	80.56	28.75
Standard deviation	15.32	36.81
Range	58.33 - 100	0 - 100

The percentage scores were analysed using a t-test. This showed that the difference in

accuracy of timing between the two schedules was highly significant (t(28) = 5.032,

p<0.001, two-tailed). Subjects in whole-interval DRO were significantly more accurate in

the timing of reinforcers than subjects in momentary DRO. Table 10.4 clearly shows the

large difference between schedules. The ranges show that subjects in whole-interval DRO

did not fall below 58.33% accuracy of timings, while subjects in momentary DRO ranged

from zero accuracy to total accuracy.

Correlations between the two measures of accuracy

As with Experiment 9, two correlations were carried out to determine whether there was a

relationship between the two measures of accuracy used in this experiment. The first

correlation was carried out between percentage error scores and percentage accuracy of

timing for whole-interval DRO. A significant negative correlation was shown (r(13) =

-0.587, p<0.05). Thus in whole-interval DRO, as the accuracy of timing reinforcers went

down, so the percentage of errors in the number of reinforcers went up. The second

correlation was carried out between percentage error scores and percentage accuracy of

timing for momentary DRO. No relationship was found between the two measures.

These results are in line with those in Experiment 9. There was a relationship between the two measures of accuracy in whole-interval DRO but there was no such relationship in momentary DRO.

Concurrent task scores

During the experiment all subjects were required to answer fifteen general knowledge questions taken from the WAIS-R test. A t-test carried out across the two schedules on subjects' WAIS-R scores showed no difference between them (mean scores on both schedules were around 10). This shows that subjects were not answering more successfully in one condition than another.

Comparisons across experiments

The following comparisons across the two experiments were carried out:

(i)   Whole-interval DRO (Experiment 9, silent condition) with whole-interval DRO (Experiment 10)

(ii)   Whole-interval DRO (Experiment 9, distracting condition) with whole-interval DRO (Experiment 10)

(iii)   Momentary DRO (Experiment 9, silent condition) with momentary DRO (Experiment 10)

(iv)   Momentary DRO (Experiment 9, distracting condition) with momentary DRO (Experiment 10)

These comparisons were carried out on the three dependent variables and are reported below.

Comparisons across experiments for difficulty of task

The first comparisons analysed subjects' reported difficulty of task, which in both experiments were scored on a six-point Likert scale. In Experiment 9, subjects in both schedules reported finding the experiment 'easy', while in Experiment 10 subjects reported

finding the experiment 'quite easy'. Unpaired two-tailed t-tests were carried out across the

two experiments and the results can be seen in Table 10.5.

Table 10.5 Statistical differences in difficulty of task across the two experiments

Comparison	t
WiDRO (Expt. 9, distraction) vs WiDRO (Expt. 10)	2.35 (p<0.05)
WiDRO (Expt. 9, silent) vs WiDRO (Expt. 10)	1.76 (p<0.1)
mDRO (Expt. 9, distraction) vs mDRO (Expt. 10)	2.03 (p<0.1)
mDRO (Expt. 9, silent) vs mDRO (Expt. 10)	1.76 (p<0.1)

There was a significant difference at the 5% level between whole-interval DRO in the

distracting condition (Experiment 9) and whole-interval DRO (Experiment 10). That is,

subjects in that condition in Experiment 9 found the experience significantly easier than

subjects in Experiment 10. The other comparisons all reached significance at the 10%

level, demonstrating a consistent trend: subjects in Experiment 9 reported their task as

being easier than subjects in Experiment 10, regardless of schedule.

Comparison across both experiments of accuracy of number of reinforcers

The percentage errors across the two experiments were analysed in pairs of t-tests (using

the same comparisons as for difficulty of task, above). There were no significant

differences between any of the comparisons, although the two whole-interval DRO

comparisons did show a trend with both approaching a 10% level of significance. That is,

more errors were made in whole-interval DRO in Experiment 10 than either of the whole-

interval DRO conditions from Experiment 9. This can be seen in the mean percentage error

scores for both experiments which are reported in Table 10.6.

Table 10.6 Mean percentage errors in number of reinforcers across both experiments

	Whole-interval DRO	Momentary DRO
Experiment 9 (Silent)	0.56	4.58
Experiment 9 (Distraction)	0.56	2.08
Experiment 10 (Concurrent task)	2.78	3.75

Comparison across both experiments of accuracy of timing of reinforcers

Although both experiments found that whole-interval DRO was timed with considerably more accuracy than momentary DRO, there were nonetheless differences between the two experiments.

Four Chi-square tests were carried out to compare the number of subjects across experiments who gave the reinforcers within the acceptable time-limit. The tests compared the number of subjects who achieved above 50% accuracy and those who achieved below 50% accuracy. There were no differences at all across whole-interval DRO, as both experiments garnered very similar results. In whole-interval DRO (Experiment 9, distraction condition), whole-interval DRO (Experiment 9, silent condition) and whole-interval DRO (Experiment 10), 100% of subjects timed reinforcers with 50% and upwards of accuracy. A significant difference was found, however, between momentary DRO (Experiment 9, distraction condition) and momentary DRO (Experiment 10) $(X^2 (1, N=30) = 6.652, p<0.01)$. That is, significantly greater numbers of subjects timed reinforcers with more than 50% accuracy in momentary DRO (Experiment 9, distraction condition), than in momentary DRO (Experiment 10). There was no such difference shown between momentary DRO (Experiment 9, silent condition) and momentary DRO (Experiment 10), although there was a slight trend in the same direction. To sum up, the number of subjects timing reinforcers accurately did not differ across experiments in whole-interval DRO, but significantly more subjects in momentary DRO (Experiment 9, distraction condition) were accurate than in momentary DRO (Experiment 10).

Subjects' percentage accuracy of timing scores were analysed across experiments using two-tailed t-tests. The results can be seen in Table 10.7.

Table 10.7 T-test results for experiment comparisons of percentage of reinforcers accurately timed

Comparison	t
WiDRO (Expt. 9, distraction) vs WiDRO (Expt. 10)	2.52 (p<0.05)
WiDRO (Expt. 9, silent) vs WiDRO (Expt. 10)	2.70 (p<0.05)
mDRO (Expt. 9, distraction) vs mDRO (Expt. 10)	2.28 (p<0.05)
mDRO (Expt. 9, silent) vs mDRO (Expt. 10)	1.80 (p<0.1)

These tests showed that in both schedules reinforcers were timed with significantly less accuracy in Experiment 10 than Experiment 9 (regardless of condition in Experiment 9). The difference was slightly less marked between momentary DRO (Experiment 9, silent condition) and momentary DRO (Experiment 10), but still reached significance at the 10% level.

That is, in all cases subjects timed reinforcers less accurately in the second experiment than in the first. To illustrate this, the mean accuracy scores for both experiments are presented in Figure 10.2 below.

Two further comparisons were made. Momentary DRO (Experiment 9, silent condition) was compared to whole-interval DRO (Experiment 10); and momentary DRO (Experiment 9, distraction condition) was also compared to whole-interval DRO (Experiment 10). Subjects in whole-interval DRO (Experiment 10) were significantly more accurate than subjects in momentary DRO (Experiment 9, silent condition) (t(28) = -2.261, p<0.05). However, there was no difference between subjects in whole-interval DRO (Experiment 10) and subjects in momentary DRO (Experiment 9, distraction condition).
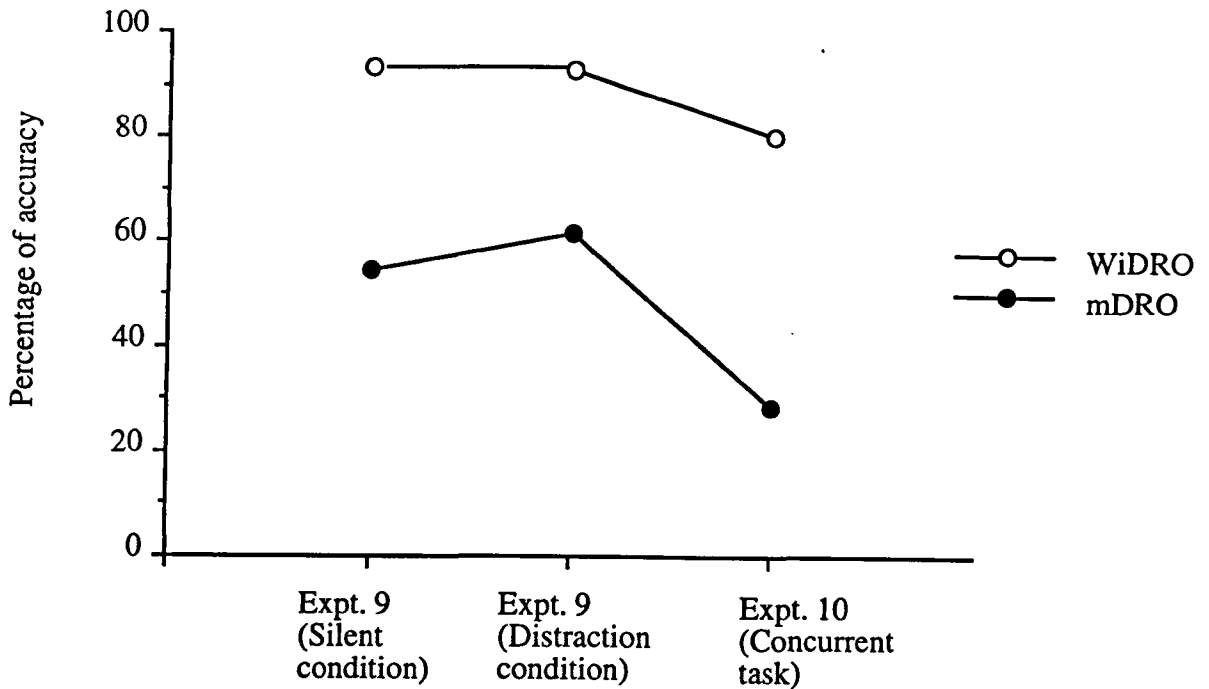
Figure 10.2 : Comparison of percentage accuracy of timing of reinforcers across the two experiments.

DISCUSSION

It was hypothesised that momentary DRO would be easier to use than whole-interval DRO when subjects are required to carry out a concurrent task. This was not supported by the data: both schedules were perceived as being equally easy to use.

It was also hypothesised that when subjects are required to carry out a concurrent task, momentary DRO will be implemented as accurately as whole-interval DRO. There was some support for this hypothesis: there was no difference between schedules in the percentage of errors made in number of reinforcers. However, on the other measure of accuracy - the timings of the reinforcers - the hypothesis was not supported. Here, subjects timed the reinforcers in whole-interval DRO with significantly greater accuracy than in momentary DRO.

The hypothesis that greater numbers of reinforcers would be delivered in momentary DRO than whole-interval DRO was supported in this experiment. These results will now be considered in more detail.

Difficulty of task

Subjects reported both schedules to be equally easy, as in Experiment 9. However, in Experiment 9, subjects in both schedules reported finding the experiment 'easy', while in the present experiment subjects reported finding the experiment 'quite easy'. Overall, then, subjects taking part in Experiment 10 found their task harder than subjects in Experiment 9. This suggests that the concurrent task in Experiment 10 made the main task more difficult; more so than the external distractions of Experiment 9, but still not enough for subjects to report the task as difficult. Subjects in Experiment 10 still found the task 'quite easy' overall.

It is possible that, as some subjects said, the task was not very distracting. Subjects were mostly able to answer questions quickly and efficiently and turn straight back to the task without missing very much of the video. Although some subjects did report some difficulty in carrying out the task at the same time as monitoring 'Danny', these comments were spread equally across schedules.

Future experiments might consider asking subjects to compare both schedules, rather than requiring them to rate the difficulty of just one schedule in isolation, as in Experiments 9 and 10. A direct comparison would provide useful information regarding whether one schedule actually was generally considered to be easier than the other.

Accuracy of the number of reinforcers

As in Experiment 9, there was a large difference between the number of reinforcers given in the two schedules, providing further evidence that this is one of the fundamental differences between momentary DRO and whole-interval DRO.

A slightly higher percentage of errors was made overall in the number of reinforcers given in momentary DRO as compared to the percentage of errors in whole-interval DRO. However this difference was not significant. In other words, subjects in whole-interval DRO made a similar number of errors to subjects in momentary DRO.

In Experiment 9, significantly higher percentage errors were made in momentary DRO. This shows that the concurrent task in Experiment 10 had an effect on the accuracy of the number of reinforcers. The means show that the task had little effect on performance in momentary DRO, but had a deleterious effect on performance in whole-interval DRO. Five times as many errors were made in whole-interval DRO in Experiment 10 than in Experiment 9, whilst the results in momentary DRO in Experiment 10 did not differ greatly from those in Experiment 9. Slightly more subjects gave the correct number of reinforcers in momentary DRO in Experiment 10 (40%) than Experiment 9 (33%). In whole-interval DRO the overall percentage of accurate subjects was considerably lower in Experiment 10 (73.33%) than Experiment 9 (93%). This result suggests that performance on momentary DRO is relatively stable regardless of concurrent task demands, whilst performance on whole-interval DRO declines in proportion to increased task demands. An experiment could be carried out to determine if there is a point at which the number or complexity of concurrent tasks causes performance on whole-interval DRO to become significantly poorer than that on momentary DRO. If it transpires that the two schedules do differ in their accuracy under differing conditions, then it would be more appropriate to use whole-interval DRO in situations where there are no other task demands and use momentary DRO where task demands are high.

Despite these differences and similarities in percentage error across momentary DRO and whole-interval DRO, it must be noted that subjects in both schedules showed impressive levels of accuracy with regard to the number of reinforcers. Percentage error on number of reinforcers was very low across both experiments and both schedules. The highest percentage error (in Experiment 9, momentary DRO silent condition) was less than 5%.

This suggests that on this measure at least, both schedules are fairly robust: subjects do not give wildly inaccurate numbers of reinforcers. No subject in either schedule gave more than two too many reinforcers, and no subject gave less than one too few. Subjects using momentary DRO were more likely to give too many reinforcers than too few. Across both experiments in momentary DRO, 42% of subjects gave too many reinforcers while 9% gave too few. In whole-interval DRO 4% of subjects gave too many reinforcers and 9% gave too few. If it is deemed important in a treatment programme to ensure that large numbers of reinforcers are available to a client, the higher density of reinforcement in momentary DRO, coupled with its apparent tendency to provide too many reinforcers rather than too few, might make it the schedule of choice.

Accuracy of timings of reinforcers

Although there was no significant difference between schedules on the measure of number of reinforcers, there was a difference regarding the accuracy of timings. As with Experiment 9, subjects in whole-interval DRO timed reinforcers with far greater accuracy than subjects in momentary DRO. All subjects in whole-interval DRO gave reinforcers with above 50% accuracy. Only 20% of subjects in momentary DRO achieved this. Large differences were also found between subjects' percentage accuracy scores, with the average score in whole-interval DRO almost three times as high as that in momentary DRO.

It would appear, therefore, that whole-interval DRO - at least on this measure of accuracy - does as well when compared to momentary DRO as in Experiment 9. However, some of the cross-experiment comparisons merit a closer look.

The number of subjects who achieved above 50% accuracy in timing reinforcers did not differ in whole-interval DRO across the two experiments - they all achieved it. This suggests that there were no differences in accuracy of timing at all for whole-interval DRO. However, the t-test comparisons which consider percentage accuracy scores contradict this. When both whole-interval DRO schedules in Experiment 9 were compared to whole-

interval DRO in Experiment 10, it was shown that subjects were significantly more accurate in the first experiment than the second. Although in Experiment 10 mean scores were still very high, particularly when compared to momentary DRO, they were significantly lower than the scores of 90% and above achieved by whole-interval DRO subjects in Experiment 9. This again demonstrates the deleterious effect that the concurrent task had on performance in whole-interval DRO. Subjects were still performing very well - as has already been noted, all subjects in Experiment 10 achieved above 50% accuracy - but when individual percentage of accuracy scores are considered it is clear that the task had a detrimental effect on performance.

Similarly, subjects in both momentary DRO conditions in Experiment 9 showed higher percentage accuracy scores than subjects in momentary DRO in Experiment 10. However, subjects' performance on error of number of reinforcers was similar in both experiments. It has already been demonstrated that there was not a relationship between accuracy in timing and number of reinforcers in momentary DRO. It is possible that in Experiment 10, the concurrent task fulfilled a similar kind of alertness function for subjects in momentary DRO as had the external interruptions of Experiment 9. However, the constant interruption of the task meant that it was easy for subjects to make errors when timing reinforcers, so that, as has been previously shown, most or all subsequent reinforcers were given at the wrong time. Certainly, percentage accuracy of timings in momentary DRO were adversely effected by the concurrent task. Overall accuracy was extremely low (28.75%), not just when compared to accuracy in whole-interval DRO but also when compared to accuracy in momentary DRO in Experiment 9.

The accuracy of momentary DRO in Experiment 9 was compared with whole-interval DRO in Experiment 10. If, as the results so far suggest, performance using whole-interval DRO was negatively effected by the concurrent task, while momentary DRO was less effected, there should be a smaller gap between the schedules across experiments than between them in Experiment 9. This was partly true: there was no difference on percentage accuracy of

timing between momentary DRO (Experiment 9, distraction condition) and whole-interval DRO (Experiment 10). The difference in accuracy, it seems, was reduced by the introduction of a concurrent task. This finding, however, was not repeated in the comparison of momentary DRO (Experiment 9, silent condition) and whole-interval DRO (Experiment 10) - here there was a significant difference between the two schedules.

Conclusion

Both schedules in the present experiment were considered by subjects to be 'quite easy'. This showed that the introduction of a concurrent task had the effect of making both schedules slightly more difficult than in Experiment 9. The concurrent task had a more detrimental effect on accuracy in whole-interval DRO than momentary DRO. Far more mistakes were made in whole-interval DRO over the number of reinforcers than in Experiment 9, while momentary DRO was not effected on this measure of accuracy. In the present experiment, equal numbers of mistakes were made in both schedules.

Accuracy of timing of reinforcers in both schedules was poorer than in Experiment 9. Performance in whole-interval DRO in Experiment 10 compared to Experiment 9 was proportionately poorer than the same comparison for momentary DRO. However, accuracy was still much higher in whole-interval DRO than momentary DRO. The lower accuracy in momentary DRO suggests that in situations where one concurrent task is required, whole-interval DRO should be the schedule of choice (assuming schedules are equally effective). However, future experiments should determine if there is a point at which the number of tasks means that momentary DRO provides greater accuracy than whole-interval DRO.

GENERAL DISCUSSION

EXPERIMENTS 9 & 10

To sum up the findings of these two experiments:

No evidence was found to support the hypothesis that momentary DRO is easier to use than whole-interval DRO: subjects found both schedules equally easy. The hypothesis that momentary DRO is more accurate than whole-interval DRO was also not supported: momentary DRO was generally less accurate than whole-interval DRO. The hypothesis that momentary DRO is more resilient than whole-interval DRO under conditions of noise and disruption was not supported. Whole-interval DRO was unaffected by such distractions, while momentary DRO was inaccurate under conditions of silence and noise. Only under conditions where a second task was introduced, and only on one measure of accuracy, was performance similar in momentary DRO and whole-interval DRO. However, momentary DRO was not at any point more accurate than whole-interval DRO. The hypothesis that greater levels of reinforcement are delivered in momentary DRO than whole-interval DRO was supported in both experiments.

The remainder of this Discussion considers some issues that should be explored in further studies of this kind.

The nature of accuracy

These experiments employed two different measures of accuracy: the number of reinforcers given and the times at which they were given. For whole-interval DRO these two measures reflected similar things; in momentary DRO they were unrelated. Accuracy of timing, in particular, was quite a sensitive measure which yielded some useful information about the differing nature of the schedules. Using such a fine measure, though, raises issues about the nature of accuracy. Many subjects whose timing was not in line with the standard set of timings were nonetheless doubtless giving reinforcers appropriately according to their

own timings: they were simply following a different time-scale. Future experiments might develop some means of investigating this, perhaps by video-taping subjects' responses. In this way, subjects' adherence to schedule rules could be determined, as well as their accuracy as measured by standard criteria.

Many subjects who were marked as being outside an acceptable time limit were only just outside it - sometimes by less than half a second. This meant that some subjects in momentary DRO whose timings were only slightly awry were classed as 'inaccurate' for all the other timings which followed, though these too may have been only fractions of seconds out. Such subjects would be accurate within their own timings, although they deviated from the standard timings. This was more of a problem in momentary DRO than whole-interval DRO because of the 'knock on' effect of inaccurate timings, as discussed previously. Again, a record of whether subjects have carried out the schedules appropriately regardless of a lack of adherence to a standard set of timings, would clarify this issue.

It remains to be seen whether either measure of accuracy is necessary for the effective management of these schedules. It is possible that as long as a reinforcer is not given when a subject is actively engaged in the inappropriate behaviour, that accuracy might not matter as much as expected. With DRO, all behaviours other than the target behaviour are reinforced - there is not a direct relationship between a single behaviour and a reinforcer. Delivering a reinforcer at a particular time in a DRO schedule may therefore not be essential; and other factors may be considerably more important in determining the effectiveness of these schedules, such as, for example, density of reinforcement.

It is also possible that there is a difference in the importance of accuracy between whole-interval DRO and momentary DRO. It is a requirement of whole-interval DRO that reinforcers are delivered after a certain period has elapsed which is free of a targeted behaviour. If a reinforcer is given too early, then the behaviour may not be absent for the

whole period, in which case the schedule rule is violated. If the reinforcer is given too late, the behaviour may have recommenced; at any rate, the connection between the interval and the reinforcer is weakened. In momentary DRO, however, the relationship is between a moment and a reinforcer, not between an interval and a reinforcer. If the reinforcer is given too early or too late, there are three possible ways that this could work to the detriment of the schedule. Obviously, if a late reinforcer coincided with the target behaviour this would be detrimental; it would be detrimental too if a late or early reinforcer reduced the number of reinforcers the client could receive. Thirdly, if the programmer was trying to implement an escalating interval schedule, then giving reinforcers at the wrong time would mean that the correct intervals were not being adhered to.

In situations other than these, accuracy may be less important. The relationship between contingency and reinforcer in momentary DRO is a relationship between a momentary absence of behaviour and the reinforcer. As long as one does not deviate too far from the correct moment, making the interval much too long or much too short, then precision accuracy is less important than it is in whole-interval DRO, when the relationship is clearly between the whole interval and the reinforcer. Most experiments in this thesis have shown that momentary DRO can work as well, and sometimes better, than whole-interval DRO. This indicates that the poorer accuracy of momentary DRO does not prevent its successful reduction of behaviour. However, future experiments could explore more thoroughly the impact of accuracy on the relative effectiveness of momentary DRO and whole-interval DRO.

Simulation and subjects

These experiments aimed to simulate a 'real-life' setting, and to a large extent successfully demonstrated the way 'naive therapists' might approach the use of DRO schedules. There were some limitations of the simulated situation, however, and these are considered below.

Although the actor in the video based his actions on a real client using real timings, his behaviour was not effected by the actions of the subjects (though some subjects seemed to think that it was. Several reported noticing his behaviour improve the more reinforcers they gave!). It would be possible, with interactive technology such as a CD ROM, to make use of a computer which allowed subjects to interact with and influence the behaviour of an actor on a film.

Although the distractions in Experiment 9 and the concurrent task in Experiment 10 were to some extent analogous with the demands of a care environment, there are improvements that could be made to future experiments. For example, distractions should be arranged so that the subjects are required to actively engage in them (the confederate could perhaps ask them some questions). A concurrent task should be more demanding, perhaps by requiring subjects to answer multiple-choice questions, or asking them to move to another desk to perform a task.

The student subjects who participated in these experiments may have had different motivations than an average care-worker. Subjects were invariably very enthusiastic; they tended to try extremely hard with the tasks and this was most evident when task demands were highest in the second experiment. It was not uncommon to see students racing from one side of the desk to the other in an attempt to answer the question quickly and get back to the video and stop-watch. Most subjects tried to carry out both tasks perfectly, even though it was made clear at the beginning that it was difficult and that the questions were intended to distract them. Care-staff may be unwilling or unable to dedicate themselves to running schedules with equal commitment. A replication of Experiment 10 could be carried out with care-staff, as this would provide useful information concerning the behaviour of people who would be required to implement such schedules on a daily basis.

Acceptability of treatments

Several subjects who were in the momentary DRO conditions reported that they didn't understand why they only had to reinforce 'Danny' for what he was doing at the 35th moment, and disregard his behaviour for the rest of the time. Subjects in whole-interval DRO conditions did not raise similar questions, and it seems therefore that whole-interval DRO makes intuitive sense and momentary DRO does not. It is possible that some of the ambivalent findings, particularly those where subjects reported on difficulty of task, may have been adversely effected by a lack of understanding. Future studies could perhaps give the background to momentary DRO and show subjects that it can be effective; this may help it make more sense to them.

Future directions

Many of the points raised by these experiments would benefit from further research. It would be useful to explore the effects of different levels of task demands on schedule accuracy. More importantly, it is necessary to determine if the accuracy of timing reinforcement has any effect on the efficacy of these schedules. It would be helpful in comparing the relative difficulty of the two schedules to ask people to experience both before making a judgement. It would be beneficial to run similar experiments with care-staff who would have a better idea of the practicalities of using such schedules.

Chapter 7

Discussion and Conclusions

AIMS OF THE THESIS

The three main aims of this thesis were as follows. First, to examine the effectiveness of momentary DRO under a variety of different conditions. The conditions which the experiments in this thesis explored were: (a) the type and rate of behaviour; (b) the language ability and general developmental level of the subject; and (c) the use of momentary DRO in combination with other treatments. This thesis also considered the previous work which suggested that momentary DRO was too weak to produce significant reductions in behaviour (e.g. Repp et al., 1983), and the more recent work which showed momentary DRO to be at least as effective as whole-interval DRO (Derwas & Jones, 1993).

The second aim of this thesis was to confirm the suggestion that momentary DRO provides higher rates of reinforcement than whole-interval DRO, and to explore whether there is a relationship between the effectiveness of momentary DRO and density of reinforcement.

The third aim of this thesis was to examine the ease of use of momentary DRO. It was hypothesised that it would be possible to carry out concurrent tasks with momentary DRO in a way that would not be possible with whole-interval DRO. Similarly, it was hypothesised that momentary DRO would be less effected by external noise and disruption because the schedule does not rely on intense and extended periods of concentration.

The effectiveness of momentary DRO

The first four experiments considered the effectiveness of momentary DRO in reducing different types and rates of behaviour. Experiments 1 and 2 showed that momentary DRO was at least as effective as whole-interval DRO in reducing high-rate stereotyped behaviour. Experiment 3 showed that momentary DRO could be as effective than whole-interval DRO in reducing low-rate disruptive behaviour. Experiment 4 also showed that momentary DRO could be as effective as whole-interval DRO in reducing disruption; however, for one subject, both schedules increased mean levels of disruption from baseline. These experiments questioned the conclusions of Repp et al. (1983) that momentary DRO was a weaker schedule than whole-interval DRO. These findings supported the results shown by Derwas and Jones (1993), that momentary DRO can be at least as effective as whole-interval DRO.

It had been previously suggested (e.g. Repp et al., 1983) that momentary DRO would not be effective with capable subjects who had good language skills. Experiments 3 and 4 did not support this; they showed that momentary DRO was as effective as whole-interval DRO when used with mildly disabled and non-disabled children.

Experiment 7 considered the role of language in more detail. The subjects in this experiment were non-learning disabled adults, who were given a partial rule for both momentary DRO and whole-interval DRO. Momentary DRO produced a quicker reduction in target behaviour than whole-interval DRO. This provides further evidence that momentary DRO can be effective with subjects who have language ability, and suggests that the reason why Repp et al. (1983) failed to show a reduction with momentary DRO was the use of very specific instructions, rather than the ability of the subjects.

Experiment 8 examined the role of momentary DRO as part of a treatment package. Momentary DRO was combined with reprimand, instructions and modelling to train a subject to discriminate between two stimuli and display stereotypy in the presence of just

one stimulus. This experiment showed that momentary DRO could be easily incorporated into a treatment package. Indeed, the effectiveness of momentary DRO may have been enhanced by being combined with other treatments.

To sum up: Experiments 1 to 6 in this thesis compared momentary DRO and whole-interval DRO within nine subjects. Momentary DRO was more effective than whole-interval DRO for two subjects, as effective for six subjects, and less effective for one subject. This means that whole-interval DRO was more effective than momentary DRO for just one subject - and this was the subject for whom both schedules increased the inappropriate behaviour from baseline. Experiment 7 compared the two schedules between subjects. Momentary DRO produced reductions of at least 50% between baseline and experimental phases for seven subjects. Whole-interval DRO produced similar reductions for four subjects. Whole-interval DRO produced increases in target behaviour in eight subjects, while momentary DRO produced similar increases for just two subjects. These seven experiments clearly show that momentary DRO is a useful reductive treatment and that for some subjects it is more effective than whole-interval DRO.

The density of reinforcement in momentary DRO

The second aim of this thesis was to confirm that momentary DRO provides higher rates of reinforcement than whole-interval DRO. Experiments 1, 6, 7, 9 and 10 all showed that this was the case. The difference in levels of reinforcement between the two schedules was statistically significant, except for one subject in Experiment 1. The clearest illustration of this trend was in Experiments 9 and 10, where 82 different subjects were trained to use one of the schedules. Without exception, they all delivered between 15 and 18 reinforcers in momentary DRO, and between 11 and 14 reinforcers in whole-interval DRO (in fact, only one subject delivered 14 reinforcers in whole-interval DRO - the rest delivered between 11 and 13). The reasons why higher rates of reinforcement are delivered in momentary DRO are outlined in Chapter 3.

Experiments 5 and 6 considered the role played by the density of reinforcement in more detail, by comparing momentary DRO and whole-interval DRO to a third schedule, momentary DRO/2, which operated on the same parameters as momentary DRO but provided greater levels of reinforcement. There was little difference between momentary DRO and momentary DRO/2 in Experiment 5; however, in Experiment 6 momentary DRO/2 did produce slightly greater reductions in behaviour than momentary DRO. In Experiment 5, the momentary DRO schedules produced slightly larger reductions than whole-interval DRO. In Experiment 6 both momentary DRO schedules were significantly more effective than whole-interval DRO.

Experiment 7 also considered the effect of increased reinforcement in momentary DRO. It was concluded that the higher levels of reinforcement were the main reason why target behaviour was reduced more quickly in momentary DRO. This was because subjects were given greater amounts of information regarding whether they were in the target area in momentary DRO.

Overall, the experiments in this thesis provide evidence that the greater reinforcement levels in momentary DRO make a major contribution to its effectiveness. Whether this is the only important component of momentary DRO is considered later in this chapter.

The ease of use and accuracy of momentary DRO

The third aim of this thesis was to determine whether momentary DRO is easier to use than whole-interval DRO. Experiments 9 and 10 explored this issue, and also considered whether momentary DRO was more accurate than whole-interval DRO under conditions of noise and concurrent tasks. Both experiments simulated a situation whereby a subject took the role of a care-assistant employing a DRO programme. These experiments showed that subjects judged both schedules to be equally easy. It is possible that if one of the experiments had asked subjects to compare both schedules, a difference in reported ease of

use would have been seen. However, given the results of Experiments 9 and 10, the evidence that momentary DRO is easier to use remains anecdotal.

Experiment 9 found that subjects using whole-interval DRO were unaffected by conditions of noise and disruption, and were highly accurate in terms of the number of reinforcers given and the times at which they were given. In contrast, subjects using momentary DRO were comparatively inaccurate regardless of condition. Experiment 10, which introduced a concurrent task, showed that subjects in whole-interval DRO were slightly more effected by the task than subjects in momentary DRO and that similar numbers of mistakes were made in both schedules regarding the number of reinforcers given. Nonetheless subjects using whole-interval DRO were still considerably more accurate on the measure of timing of reinforcers.

These experiments showed that momentary DRO was generally less accurate than whole-interval DRO, except under one condition on one measure, when it was as accurate. The relative importance of accuracy in the two schedules was discussed, and it was concluded that accuracy, particularly in timing of reinforcers, may be less essential in momentary DRO than whole-interval DRO. Thus, although Experiments 9 and 10 found momentary DRO to be less accurate than whole-interval DRO, this may not be a significant problem. As the other experiments in this thesis have shown, momentary DRO is a highly effective schedule. If Experiments 9 and 10 were genuine demonstrations of how momentary DRO is carried out, this suggests that accuracy - or lack of accuracy - does not interfere with its effectiveness.

The experiments in this thesis have shown momentary DRO to be a successful behaviour reduction technique. The next section considers the possible factors that contribute to its effectiveness.

Density of reinforcement

One possible reason for the success of momentary DRO - high density of reinforcement - has been explored in previous chapters. The results of Experiment 1 were particularly suggestive that greater levels of reinforcement were responsible for the differences seen between whole-interval DRO and momentary DRO. One of the subjects in this experiment did not receive significantly greater levels of reinforcement in momentary DRO, and for this subject both schedules were equally effective. For the other subject, who did receive significantly greater levels of reinforcement in momentary DRO, momentary DRO was more effective.

Five possible reasons why higher levels of reinforcement could lead to greater behaviour reduction were outlined in Chapter 3: more rapid extinction, the competing function of programmed reinforcers, satiation of reinforcement, speedier learning, and increased attention. Increased attention in particular provides an explanation for some of the results seen in this thesis, such as in Experiment 2, where it seemed the subject was responding to the attention from the experimenter rather than to the programmed reinforcer. Sometimes, therefore, attention proved to be a more potent reinforcer than the one which had been programmed. Related to this is the novel and enriched environment which increased contingent reinforcement and attention bring about. The subject may feel less ignored and less isolated; suddenly, regardless of what they are doing, they are receiving high levels of reinforcement and attention. For a person who spends little time engaged with others this must have some impact on the way they behave. As discussed in Chapter 3, the possibility that increased attention or the novelty of increased stimulation may be more potent reinforcers than the programmed reinforcers has implications for the validity of functional analysis. This is a factor which should be taken into account, particularly when running programmes in deprived environments.

The experiments in this thesis strongly suggest that density of reinforcement plays an important role in the effectiveness of momentary DRO. Future experiments could examine this factor systematically. For example, a simulation could be used, as in Experiment 7, which compared several momentary DRO schedules, each providing different levels of reinforcement. Identical interval lengths could be used in each case, but one schedule could provide reinforcement every other interval, while another could provide it every third interval. Such an experiment would contribute further information regarding the role of reinforcement in momentary DRO.

The paradox of momentary DRO

Regardless of the role played by reinforcement, the success of momentary DRO in behaviour reduction is, at first glance, a paradox. In this schedule, unlike in whole-interval DRO, the subject can display the target behaviour during intervals without jeopardising the possibility of reinforcement. It may seem surprising that under such circumstances momentary DRO should actually reduce a target behaviour - if the subject incurs no penalty for displaying the target behaviour one might expect it to increase or at least remain at baseline levels. However, this process may help explain why momentary DRO was more successful than whole-interval DRO in several of the experiments in this thesis. In whole-interval DRO, it may be some time before a behaviour comes into contact with the contingencies, particularly if the behaviour occurs at a high rate. If an individual displays the target behaviour at all, then a reinforcer is not delivered and a new interval begins. In whole-interval DRO, therefore, the opportunities to earn reinforcement and for the behaviour to come under the control of the contingencies may be quite limited. By contrast, in momentary DRO the individual can present the target behaviour and still earn reinforcement, as long as the behaviour is absent at the designated moment. This means that not only are there greater opportunities for reinforcement (as previously discussed) but that there are many more chances for the behaviour to come under contingency control. This strongly suggests that where contingency control is essential for the success of a programme (e.g. where subjects do not have language, or where rules are not provided),

momentary DRO should be the schedule of choice. In cases where subjects are provided with rules, contingency control may be weakened by rule-governed behaviour and so there may be less advantage to using momentary DRO.

The results of Experiment 7 provide support for the suggestion that behaviour comes under contingency control more quickly in momentary DRO than in whole-interval DRO.

The elements of non-contingent reinforcement and punishment

Momentary DRO is in some ways similar to a non-contingent procedure (Mazaleski et al., 1993). Both procedures provide high levels of reinforcement; and both allow the subject access to the target behaviour. The difference is that momentary DRO has a built-in control to avoid inadvertently reinforcing the target behaviour. Studies such as that by Vollmer et al. (1993) have shown that NCR can be a highly effective reductive technique. Future experiments should consider whether there are significant differences between NCR and momentary DRO outcomes and procedures, as this would help to provide more information regarding why momentary DRO is effective.

Occasionally, whole-interval DRO has been described as a negative punishment procedure (e.g. Rolider & Van Houten, 1990) as it is based on the reduction of reinforcement following a response. It is doubtful that momentary DRO can be regarded in the same way. Although (unlike non-contingent reinforcement) it requires that certain conditions are met before reinforcement is delivered, the individual is not penalised for every instance of the targeted behaviour. Several of the experiments in this thesis have shown that, even when inappropriate behaviour occurs very frequently, subjects still receive high levels of reinforcement in momentary DRO. The possibility that whole-interval DRO may be effective in part because of a punitive element does not appear from the evidence to be true of momentary DRO.

What is reinforced in DRO schedules?

There has been some debate regarding the terminology of differential reinforcement, which reflects the uncertainty regarding what is actually reinforced during DRO schedules. In the original paper on DRO, Reynolds (1961) termed it the 'differential reinforcement of other behaviour'. However, several commentators suggested that 'differential reinforcement of not responding' or 'differential reinforcement of pausing' would be more appropriate (Poling & Ryan, 1982; Zeiler, 1970, 1979). These latter terms imply that it is an *absence* of behaviour which is reinforced during DRO. This does not tally with Skinner's (1953) definition of reinforcement as something which leads to an increased probability of response occurrence. The reinforcement of non-responding suggests the opposite: that the reinforcement in DRO *decreases* a behaviour. To get round this difficulty, it has been proposed that not responding should be regarded as a response class in its own right (Poling & Ryan, 1982; Vollmer & Iwata, 1992). Proponents of this theory claim that if non-responding is regarded as a response class then it can increase, even though the behaviour to which it refers actually decreases. However, it is not clear what an increase in non-responding actually means. The implication is that the individual emits less behaviour than during baseline; but it is not clear what someone 'behaving less' actually means, nor how one could tell if this was in fact happening. This rather convoluted and inconclusive explanation is not necessary if it is accepted that what is actually being reinforced during DRO is all behaviour other than the one targeted for reduction. If this is so, it must be true that other behaviours increase as they are reinforced. The matching law (Herrnstein, 1961) states that behaviours which are associated with reinforcement increase in direct proportion to behaviours not associated with reinforcement. In DRO schedules, therefore, all behaviours other than the one targeted should increase as the targeted behaviour decreases. Whether this does actually happen is an empirical question which could be investigated by keeping a real-time record of all behaviours. Any increases or decreases could be analysed, and in this way the effect of differential reinforcement on the rates of behaviour could be determined.

Chapter 1 outlined the reported side-effects of whole-interval DRO. The next section considers the side-effects of momentary DRO that have been noted in experiments in this thesis.

SIDE-EFFECTS OF MOMENTARY DRO

Negative side-effects

The experiments in this thesis did not demonstrate many negative side-effects of momentary DRO. For one subject (Perry in Experiment 4) the targeted behaviour increased relative to baseline under momentary DRO. However, whole-interval DRO also slightly increased this subject's target behaviour. It was concluded that Perry's behaviour had been adversely effected by extraneous events and that neither schedule was able to compete with these conflicting variables. This was the only subject whose behaviour increased under momentary DRO, therefore, which compares favourably to the three subjects (including Perry) whose targeted behaviour increased under whole-interval DRO.

Also in Experiment 4, there was evidence to support the findings of Cowdery et al. (1990) that DRO schedules can produce negative emotional side-effects. Perry became aggressive when he was told that an interval had ended and he hadn't earned a star. This problem was quickly avoided by ceasing to tell him that the interval had ended. The schedule which was first introduced for Perry was whole-interval DRO, and so by the time momentary DRO was introduced the instructions had already been changed to avoid informing him of the interval end. Therefore it is not known if a similar aggressive reaction would have occurred under momentary DRO. Vollmer et al. (1993) suggested that the side-effects shown by Cowdery et al. (1990) were due to extinction, but the pattern of results in Experiment 4 suggests that the reason for the aggression shown here was more likely to be due to a punishment element. This was the only case of negative emotion during the experiments in this thesis, which may be because it was the only experiment in which subjects were informed if they had not earned a reinforcer. It is possible therefore that such

instructions introduce negative side-effects into whole-interval DRO of the type more usually associated with punishment techniques. Whether such instructions can also produce side-effects in momentary DRO is not yet known.

Effect of schedules on monitored behaviours

It might be anticipated that if a DRO schedule suppressed a particular inappropriate behaviour, then others might increase to replace it. In order to assess this, many of the experiments in this thesis took measures of other inappropriate behaviours. Increases occurred only rarely. In Experiment 1, Matthew's monitored behaviour showed a slight increase under whole-interval DRO, and in Experiment 4, Mark's monitored behaviour increased under both schedules. Mostly, however, both DRO schedules suppressed non-targeted inappropriate behaviour. One possible reason for this is that inappropriate behaviours were often presented together, so that when one was reduced the other also occurred less frequently. Some baseline measures showed that behaviours did sometimes follow similar patterns to each other. However, other baselines showed that the inappropriate behaviours did not always occur together. Another reason for suppression of monitored behaviours may be provided by a homeostatic explanation (e.g. Goodall & Corbett, 1982). An individual's various inappropriate behaviours might provide an overall level of reinforcement. If additional reinforcement is provided as reward for the absence of a targeted inappropriate behaviour, this might provide enough overall levels of reinforcement to render unnecessary the other inappropriate behaviours. On one occasion the schedules suppressed the monitored behaviour even more effectively than the targeted behaviour (Paddy in Experiment 2). Functional analyses could be carried out for the monitored behaviours as well as the targeted one, to allow some prediction of the effects of the reinforcers on all recorded behaviours.

Inadvertent strengthening of other inappropriate behaviours did not appear to take place, presumably because reinforcement was never provided if any other undesirable behaviours were occurring.

Overall, both DRO schedules were fairly effective in suppressing monitored behaviour, although momentary DRO was more successful. Momentary DRO produced reductions in monitored behaviour in five subjects, and whole-interval DRO in two. It thus appears that this is a reasonably consistent side-effect of momentary DRO and one which deserves further attention. It is important that future studies take measurements of non-targeted behaviour so that the extent of this phenomenon can be examined.

Experiment 4 monitored a desirable behaviour: engagement in classroom tasks. This experiment showed that both schedules could produce increases in on-task behaviour, although there were differences between the two subjects. The mean of Perry's on-task behaviour was higher during both interventions than baseline. However, Mark's on-task behaviour fell slightly during momentary DRO, and increased again under whole-interval DRO. Future studies should consider monitoring desirable behaviours as well as undesirable ones, in order to assess the full impact of DRO schedules on behaviour.

In conclusion: momentary DRO, like whole-interval DRO, appears to have few negative side-effects; those which have been noted such as negative emotions and increases in monitored behaviour occur only rarely. The most frequently occurring side-effect was the positive one of suppression of other inappropriate behaviours.

PRACTICAL IMPLICATIONS

Ease of use

The experiments in this thesis all support the use of momentary DRO for reducing inappropriate behaviour. Under a variety of different circumstances it has been shown to be at least as effective as whole-interval DRO, and is more effective than whole-interval DRO in suppressing other inappropriate behaviours not directly targeted by the schedule. It is predicted, though not yet confirmed, that people who worked with both schedules would

find momentary DRO easier to use; all that can be said with certainty at the moment is that momentary DRO is seen to be as easy as whole-interval DRO. Certainly it is the experience of the author that momentary DRO is far easier to use in an applied setting than whole-interval DRO. It is not always possible, nor desirable, to work with clients in a separate room; often, it is necessary to carry out treatments in busy, noisy environments with many other clients and staff present. Using momentary DRO it is possible to hold conversations with staff and clients, carry out other tasks, and fulfil the requirements of the schedule in a way that is not possible with whole-interval DRO.

Resetting the stop-watch after each interval

All the experiments in this thesis employed a momentary DRO procedure which required the stop-watch to be stopped at the end of an interval, and the watch to be reset in order to begin the next interval. An alternative procedure would be to use a signalling device, such as a pre-recorded tape, in place of a stop-watch (e.g. Barton et al., 1986). The end of intervals could thus be signalled automatically, and the programmer would just have to look up at the signal and make a reinforcement decision. This procedure was not used in this thesis for two reasons. Firstly, because stopping and resetting at the end of intervals is widely used in whole-interval DRO procedures, and the experiments in this thesis were designed to compare the two schedules as fairly as possible. Secondly, because a difficulty arises if continuous intervals are employed; with no break between intervals, tangible reinforcers will coincide with the next interval. That is, the interval after a reinforcer has been delivered will contain within it the subject receiving and attending to the reinforcer. This may produce an unclear picture of the effect of the schedule on the behaviour. The schedules employed in this thesis all contained a brief pause when a reinforcer was delivered, so that the reinforcer would not coincide with the next interval. In this way, it is believed that a more accurate account of the effect of the schedules on behaviour was obtained. However, sometimes tokens are used as reinforcers, and these would be less likely to interfere with the running of the schedule. In such a situation, the programmer

could use continuous intervals and exchange the tokens for primary reinforcers at the end of the session. This may enhance the ease of implementation of momentary DRO.

Accuracy

Momentary DRO was found to be less accurate than whole-interval DRO, though the gap between the two schedules appeared to close as a concurrent task was introduced. However, accuracy may be less important to the effective running of momentary DRO than it is for whole-interval DRO. Paradoxically, the poorer levels of accuracy seen in momentary DRO may be one of its strengths; it appears that momentary DRO can still work extremely well despite inaccuracies in timing the reinforcers. Therefore, the programmer using momentary DRO need not worry if small mistakes are made (as long as a reinforcer doesn't inadvertently coincide with the target behaviour or any other undesirable behaviour), for this should not prevent momentary DRO from working.

Speed of reduction

There are practical implications concerning the suggestion made earlier in this chapter that momentary DRO allows more opportunities for contingency control than whole-interval DRO. If momentary DRO therefore produces quicker control than whole-interval DRO, it should be the schedule of choice where a speedy reduction is important. There was some evidence in this thesis that momentary DRO worked more rapidly; in Experiment 7 this was clearly the case. In cases where the use of rules is not possible, and where a quick reduction is necessary, momentary DRO may have practical advantages over whole-interval DRO.

Strength of reduction and generalisability

Neither DRO schedules produced total and consistent reductions in target behaviour, a fact which has been noted before (Barton et al., 1986). There are several possible reasons for this, including treatment decay, reinforcer satiation and extinction bursts. It is also possible that a behaviour which has been presented at very high levels for many months or years is too deeply ingrained to be totally eliminated unless it is replaced by another behaviour which serves the same function. The distraction of a DRO schedule may be temporarily enough to reduce the behaviour substantially, but such an effect will probably not last in the absence of the treatment. This thesis did not produce any firm conclusions about generalisability. However, it is recommended that future DRO programmes contain one or more of the following elements:

(i)   Once the target behaviour has been reduced, the results of the functional analysis are used to determine a more appropriate behaviour which can be taught to the client as a replacement behaviour. A period of training and intense reinforcement should be carried out to establish it firmly in the client's repertoire.

(ii)   Once the target behaviour has been reduced, appropriate natural reinforcers such as attention from carers should be phased in, so that by the time the treatment is removed, the behaviour is under the control of contingencies present in the environment.

(iii)   If generalisability to other settings is desired, this should be programmed into the treatment.

(iv)   An escalating momentary DRO schedule should be used, so that intervals are gradually increased. If possible, the treatment should ultimately be faded out and once again, natural reinforcement slowly introduced.

The use of momentary DRO with other inappropriate behaviours

Because of the ethical problems surrounding testing new treatments on more dangerous behaviours, none of the experiments in this thesis were carried out with inappropriate behaviours such as self-injury or aggression. However, it is proposed that momentary DRO will be at least as effective as whole-interval DRO in working with these behaviours.

Whole-interval DRO has been shown to be effective in reducing a wide range of inappropriate behaviours, including self-injurious behaviour (Anderson et al., 1978; Cowdery et al., 1990; Matson & Keys, 1990; Vollmer et al., 1993), aggression (Andrews, 1988; Redmon, 1987; Whitaker, 1992), inappropriate masturbation (Foxx et al., 1986), stereotyped behaviour (Barton et al., 1986; Haring et al., 1986; Kennedy & Haring, 1993), and disruptive behaviour (Poling et al., 1978; Repp et al., 1976). The findings in this thesis suggest that momentary DRO will be at least as effective as whole-interval DRO in reducing these behaviours; future research will be necessary to determine whether this is the case. It is of course important that standard practices of functional analysis are carried out before momentary DRO is employed with any behaviour.

FUTURE DIRECTIONS

This thesis has shown that momentary DRO is a positive reduction treatment worth considering. It is at least as effective as the more widely used whole-interval DRO, and it provides higher levels of reinforcement. It has no more negative side-effects than whole-interval DRO and has similar positive side-effects. However, a great deal remains unknown, and therefore some further experiments and analyses are proposed.

It is important to determine if momentary DRO is effective in reducing inappropriate behaviours other than stereotypy and disruption. The findings from these experiments strongly suggest that momentary DRO *will* be effective with other behaviours; however, this needs to be clarified. It is also important to determine if momentary DRO can exert control over behaviours which occur at a lower rate than those examined in Experiments 3 and 4. The role of instructions and signals in momentary DRO should be clarified and the level of instructions which allow this schedule to operate most effectively should be determined. For example, an experiment could be devised whereby non-learning disabled subjects are provided with different amounts of information regarding the momentary DRO schedule. One group of subjects could be given full and explicit instructions; a second

group could be given instructions which explain only that reinforcers will be provided for an absence of the target behaviour, without specifying details of the momentary nature of the schedule; and a third group could be given no instructions at all. Such an experiment would provide useful information regarding the effect of rules on the schedule, and whether partial rules are more effective than full rules, or whether no rules are more effective still.

It is not yet fully understood why momentary DRO can be more effective than whole-interval DRO, or produce a speedier reduction. The findings of some of the experiments in this thesis provide strong evidence that the higher levels of reinforcement enhance the speed of learning, but this needs to be clarified experimentally. There also may be a role played by the greater opportunities for the behaviour to contact the contingencies: this needs further exploration.

Vollmer et al. (1993) suggested that non-contingent reinforcement was a viable alternative to whole-interval DRO. It has many features in common with momentary DRO, and it would provide even higher levels of reinforcement and should be extremely easy to implement. However, it has the potential problem of inadvertent conditioning. Future experiments could explore the similarities and differences between momentary DRO and non-contingent reinforcement. This should provide some useful information with regard to how such treatments bring about reductions in behaviour, including the role played by high levels of reinforcement, access to the target behaviour and contingency control, and the changes in the environment brought about by high levels of reinforcement and attention.

The findings regarding ease of use of momentary DRO do not support the anecdotal evidence that it is easier to implement than whole-interval DRO. A further experiment should be carried out to clarify this. The most appropriate means of determining the relative ease of the two schedules would be to ask care-staff to carry out both schedules in a simulated and an applied setting. Ease of use could be determined by a combination of

observation and interview. It is proposed that this might offer a more complete answer to the question of whether one schedule is easier to use than another, as subjects would be able to compare the schedules and assess their relative ease under different settings. Other opinions of care-staff, whether for example they would be happy to use momentary DRO, should also be sought. It has been noted that a programme is more likely to succeed if the care-staff approve of it (Burgio, Whitman & Reid, 1983) and so attitudes to momentary DRO would provide useful information on its likely future use.

CONCLUSIONS

This thesis has shown that momentary DRO was dismissed prematurely when Repp et al. (1983) suggested it was only suitable for maintaining reductions begun by whole-interval DRO. The experiments in this thesis support the findings of Derwas and Jones (1993): momentary DRO is strong enough to produce substantial reductions in behaviour and in some cases is more effective than whole-interval DRO.

This thesis found that momentary DRO was as effective as whole-interval DRO in reducing both high-rate stereotypy and low rate disruptive behaviour. For a number of subjects, momentary DRO was more effective than whole-interval DRO. Momentary DRO was shown to be useful as part of a treatment package, and it was suggested that its effectiveness might sometimes be enhanced by being used in conjunction with other treatments.

Contrary to reports by Repp et al. (1983) that momentary DRO would be ineffective for capable people with well developed language skills, experiments in this thesis have shown that momentary DRO can be used successfully with non-learning disabled children and adults. Indeed, momentary DRO produced a faster reduction than whole-interval DRO amongst non-learning disabled adults. The role played by instructions (and the lack of instructions) in these studies was not clear. It is possible that if full instructions are provided and understood, momentary DRO may not operate successfully. However, this has yet to be determined.

This thesis confirmed that momentary DRO allows higher levels of reinforcement to be provided, and evidence from some experiments suggests that this may be the most important factor in its success. This requires further empirical confirmation, however, as does the related suggestion that momentary DRO allows greater opportunities for behaviour to contact the contingencies.

Contrary to predictions, momentary DRO was not rated by subjects as being easier to use than whole-interval DRO; however subjects were asked to use only one schedule and rate it after use. It is proposed that if subjects had been asked to compare both schedules, momentary DRO would have been rated as the easier to use. Momentary DRO was found to be less accurate than whole-interval DRO, though under conditions where a concurrent task was introduced, the gap in accuracy between the two schedules closed somewhat. However, overall whole-interval DRO was considerably more accurate. It was suggested that accuracy was actually less important in momentary DRO than in whole-interval DRO, as its effective implementation does not rely on the relationship between the complete absence of a behaviour and a reinforcer.

Overall, the experiments in this thesis confirm that momentary DRO is at least as effective a reductive technique as whole-interval DRO. As it provides higher levels of reinforcement and may be easier to use, it warrants further examination and attention than it has hitherto received. The evidence from this thesis suggests that momentary DRO should be considered as a useful and effective tool in behaviour reduction programmes.

REFERENCES

Adlem, A., & MacDonald, D. (1992). Covariation of stereotyped behaviour and engagement. *Behaviour Change, 9,* 8-13.

Ager, A. (1991). Effecting sustainable change in client behaviour: The role of the behavioural analysis of service environments. In B. Remington (Ed.), *The challenge of severe mental handicap.* Chichester: John Wiley & Sons Ltd.

Allen, K.E., & Harris, F.R. (1966). Elimination of a child's scratching by training the mother in reinforcement procedures. *Behaviour Research and Therapy, 4,* 79-84.

Anderson, L., Dancis, J., & Alpert, M. (1978). Behavioral contingencies and self-mutilation in Lesch-Nyan disease. *Journal of Consulting and Clinical Psychology, 46,* 592-536.

Andrews, M. (1988). Changing disruptive behaviour in an Adult Training Centre client. *Behavioural Psychotherapy, 16,* 108-114.

Azrin, N.H., Gottlieb, L., Wesolowski, M.D., & Rahn, T. (1975). Eliminating self-injurious behavior by educative procedures. *Behavior Research and Therapy, 13,* 101-110.

Azrin, N.H., Kaplan, S.J., & Foxx, R.M. (1973). Autism reversal: Eliminating stereotyped self-stimulation of retarded individuals. *American Journal of Mental Deficiency, 78,* 241-248.

Azrin, N.H., & Wesolowski, M.D. (1980). A reinforcement plus interruption method of eliminating behavioral stereotypy of profoundly retarded persons. *Behavior Research and Therapy, 18,* 113-119.

Bandura, A. (1965). Influence of model's reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology, 1,* 589-595.

Barrett, R.P., Matson, J.L., Shapiro, E.S., & Ollendick, T.H. (1981). A comparison of punishment and DRO procedures for treating stereotypic behavior of mentally retarded children. *Applied Research in Mental Retardation, 2,* 247-256.

Barton, E.S., & Barton, J.L. (1978). A case report on the use of DRO in the treatment of smoking. *Addictive Behaviors, 3,* 1-4.

Barton, L.E., Brulle, A.R., & Repp, A.C. (1986). Maintenance of therapeutic change by Momentary DRO. *Journal of Applied Behavior Analysis, 19,* 277-282.

Berg, W.K., & Wacker, D.P. (1991). The assessment and evaluation of reinforcers for individuals with severe mental handicap. In B. Remington (Ed.), *The challenge of severe mental handicap.* Chichester: John Wiley & Sons Ltd.

Bijou, S.W. (1976). *Child development III: Basic state of early childhood.* Englewood Cliffs, NJ: Prentice Hall.

Borreson, P.M. (1980). The elimination of a self-injurious avoidance response through a forced running consequence. *Mental Retardation, 18,* 73-77.

Bostow, D.E., & Bailey, J.B. (1969). Modifications of severe disruptive and aggressive behavior using brief timeout and reinforcement procedures. *Journal of Applied Behavior Analysis, 2,* 31-37.

Brusca, R.M., Nieminen, G.S., Carter, R., & Repp, A.C. (1989). The relationship of staff contact and activity to the stereotypy of children with multiple disabilities. *Journal of the Association for Persons with Severe Handicaps, 14,* 127-136

Burgio, L., Whitman, T., & Reid, D. (1983). A participative management approach for improving direct care staff performance in an institutional setting. *Journal of Applied Behavior Analysis, 16,* 37-53.

Carr, E.G., & Durand, V.M. (1985a). The social-communicative basis of severe behavior problems in children. In S. Reiss & R.R. Bootzin (Eds.), *Theoretical issues in behavior therapy.* New York: Academic Press.

Carr, E.G., & Durand, V.M. (1985b). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18,* 111-126.

Christensen, A.P., & Sanders, M.R. (1987). Habit reversal and differential reinforcement of other behaviour in the treatment of thumb-sucking: An analysis of generalization and side-effects. *Journal of Child Psychology and Psychiatry, 28,* 281-295.

Clarke, A.D.B., & Clarke, A.M. (1987). Research on mental handicap, 1957-1987: A selective review. *Journal of Mental Deficiency Research, 31*, 317-328.

Cooper, J.O. (1987). Decreasing behavior with differential reinforcement. In J.O. Cooper, T.E. Heron & W.L. Heward *Applied behavior analysis.* Colombus: Merrill.

Corte, H.E., Wolf, M.M., & Locke, B.J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis, 4,* 201-213.

Cowdery, G.E., Iwata, B.A., & Pace, G.M. (1990). Effects and side effects of DRO as treatment for self-injurious behavior. *Journal of Applied Behavior Analysis, 23,* 497-506.

Crawford, J., Brockel, B., & Schauss, S. (1992). A comparison of methods for the functional assessment of stereotypic behavior. *Journal of the Association for Persons with Severe Handicaps, 17,* 77-86.

Cullen, C. (1983). Implications of functional analysis. *British Journal of Clinical Psychology, 22,* 137-138.

Deitz, D.E.D., & Repp, A.C. (1983). Reducing behavior through reinforcement. *Exceptional Education Quarterly, 3,* 34-46.

Deitz, S.M., & Repp, A.C. (1973). Decreasing misbehavior through the use of DRL schedules of reinforcement. *Journal of Applied Behavior Analysis, 6,* 457-463.

Deitz, S.M., & Repp, A.C. (1974). Differentially reinforcing low rates of misbehavior with normal elementary school children. *Journal of Applied Behavior Analysis, 7,* 622.

Denny, M. (1980). Reducing self-stimulatory behavior of mentally retarded persons by alternative positive practice. *American Journal of Mental Deficiency, 84,* 610-615.

Derwas, H., & Jones, R.S.P. (1993). Reducing stereotyped behavior using momentary DRO: An experimental analysis. *Behavioral Residential Treatment, 8,* 45-53.

Didden, R., Duker, P.C., & Korzilius, H. (In press). A meta-analytic study on treatment effectiveness for problem behaviors with individuals with a mental handicap. *American Journal on Mental Retardation.*

Donnellan, A.M., Mirenda, P.L., Mesraos, R.A., & Fassbender, L.L. (1984). Analyzing the communicative functions of aberrant behavior. *Journal of the Association for Persons with Severe Handicaps, 9,* 201-212.

Durand, V.M. (1982). Analysis and intervention of self-injurious behavior. *Journal of the Association for the Severely Handicapped, 7,* 44-53.

Durand, V.M. (1990). *Severe behavior problems: A functional communication training approach.* New York: Guilford Press.

Durand, V.M., & Berotti, D. (1991). Treating behavior problems with communication. *Bulletin of the American Speech-Language-Hearing Association,* November, 37-39.

Durand, V.M., & Crimmins, D.B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18,* 99-117.

Dyer, K. (1987). The competition of autistic stereotyped behavior with usual and specially-assessed reinforcers. *Journal of Autism and Developmental Disabilities, 8,* 607-626.

Egel, A.L. (1981). Reinforcer variation: Implications for motivating developmentally disabled children. *Journal of Applied Behavior Analysis, 14, 345-350.*

Favell, J.E., McGimsey, J.F., & Jones, M.L. (1980). Rapid eating in the retarded: Reduction by nonaversive procedures. *Behavior Modification, 4,* 481-492.

Favell, J.E., McGimsey, J.F., & Jones, M.L. (1978). The use of physical restraint in the treatment of self-injury and as positive reinforcement. *Journal of Applied Behavior Analysis, 11,* 225-241.

Ferster, C.B., & Skinner, B.F. (1957). *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Forehand, R., & Baumeister, A.A. (1970). The effect of auditory and visual stimulation on stereotyped rocking behavior and general activity of severe retardates. *Journal of Clinical Psychology, 26,* 426-429.

Foxx, R.M., & Azrin, N.H. (1973). The elimination of autistic self-stimulatory behavior by overcorrection. *Journal of Applied Behavior Analysis, 6,* 1-14.

Foxx, R.M., McMorrow, M.J., Fenlon, S., & Bittle, R.G. (1986). The reductive effects of reinforcement procedures on the genital stimulation and stereotypy of a mentally retarded adolescent male. *Analysis and Intervention in Developmental Disabilities, 6*, 239-248.

Frankel, F., Moss, D., Schofield, S., & Simmons, J.Q. (1976). Case study: Use of differential reinforcement to suppress self-injurious and aggressive behavior. *Psychological Reports, 39*, 843-849.

Freeman, B.J., Graham, V., & Ritvo, E.R. (1975). Reduction of self-destructive behavior by overcorrection. *Psychological Reports, 37*, 446.

Friman, P.C., Barnard, J.D., Altman, K., & Wolf, M.M. (1986). Parent and teacher use of DRO and DRI to reduce aggressive behavior. *Analysis and Intervention in Developmental Disabilities, 6*, 319-330.

Goetz, E.M., Holmberg, M.C., & LeBlanc, J.M. (1975). Differential reinforcement of other behavior and noncontingent reinforcement as control procedures during the modification of a preschooler's compliance. *Journal of Applied Behavior Analysis, 8*, 77-82.

Goodall, E., & Corbett, J. (1982). Relationships between sensory stimulation and stereotyped behaviour in severely mentally retarded autistic children. *Journal of Mental Deficiency Research, 26*, 163-175.

Green, C.W., Reid, D.H., White, L.K., Halford, R.C., Brittain, D.P., & Gardner, S.M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis, 21*, 32-43.

Groden, G. (1989). A guide for conducting a comprehensive behavioral analysis of a target behavior. *Journal of Behavior Therapy and Experimental Psychiatry, 20*, 163-170.

Haring, T.G., Breen, C.G., Pitts-Conway, V., & Gaylord-Ross, R. (1986). Use of differential reinforcement of other behavior during dyadic instruction to reduce stereotyped behavior of autistic students. *American Journal of Mental Deficiency, 90*, 694-702.

Harris, S.L., & Wolchik, S.A. (1979). Suppression of self-stimulation: Three alternative strategies. *Journal of Applied Behavior Analysis, 12*, 185-198.

Hayes, S.C., Brownstein, A.J., Haas, J.R., & Greenway, D.E. (1986). Instructions, multiple schedules, and extinction: Distinguishing rule-governed from schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior, 46,* 137-147.

Hergenhahn, B.R., & Olson, M.H. (1993). *An introduction to theories of learning.* Fourth edition. Englewood Cliffs, NJ: Prentice-Hall.

Herrnstein, R.J. (1961). Relative and absolute strengths of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4,* 267-272.

Homer, A.L., & Peterson, L. (1980). Differential reinforcement of other behavior: A preferred response elimination procedure. *Behavior Therapy, 11,* 449-471.

Horner, R.D., & Barton, E.S. (1980). Operant techniques in the analysis and modification of self-injurious behavior: A review. *Behavior Research of Severe Developmental Disabilities, 1,* 61-91.

Hoyert, M.S., & Zeiler, M.D. (1995). Stimulus properties of the differential-reinforcement-of-not-responding schedule. *The Psychological Record, 45,* 73-93.

Iwata, B.A., Dorsey, M.F., Slifer, K.J., Bauman, K.E., & Richman, G.S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities, 2,* 3-20.

Iwata, B.A., & Fisher, W. (1994). Functional analysis approaches to behavioral assessment and treatment [Special issue]. *Journal of Applied Behavior Analysis, 27* (2).

Jones, R.S.P. (1991a). Reducing inappropriate behaviour using non-aversive procedures: Evaluating differential reinforcement schedules. In B. Remington (Ed.), *The challenge of severe mental handicap.* Chichester: John Wiley & Sons Ltd.

Jones, R.S.P. (1991b). Stereotypy: Challenging behaviour or adaptive response. *Scandinavian Journal of Behaviour Therapy, 20,* 25-40.

Jones, R.S.P., & Baker, L.J.V. (1988a). Reducing stereotyped behaviour using differential reinforcement: A comparison of DRO and DRI schedules. *Mental Handicap, 16,* 171-174.

Jones, R.S.P., & Baker, L.J.V. (1988b). The differential reinforcement of incompatible responses in the reduction of self-injurious behaviour: A pilot study. *Behavioural Psychotherapy, 16,* 323-328.

Jones, R.S.P., & Baker, L.J.V. (1989). Reducing stereotyped behaviour: A component analysis of the DRI schedule. *British Journal of Clinical Psychology, 28,* 255-266.

Jones, R.S.P., Baker, L.J.V., & Murphy, M.J. (1988). Reducing stereotyped behaviour: The maintenance effects of a DRO reinforcement procedure. *Journal of Practical Approaches to Developmental Handicap, 12,* 24-30.

Jones, R.S.P., & Carter, M.F. (1991). The relationship between environmental noise and rates of stereotyped behaviour: An ecological analysis. *The Irish Journal of Psychology, 12,* 406-417.

Jones, R.S.P., Walsh, P.G., & Sturmey, P. (1995). *Stereotyped movement disorders.* Chichester: John Wiley & Sons.

Jones, R.S.P., Wint, D., & Ellis, N.C. (1990). The social effects of stereotyped behaviour. *Journal of Mental Deficiency Research, 34,* 261-268.

Kantor, J.R. (1959). *Interbehavioral psychology.* Granville: Principia Press.

Kazdin, A.E. (1982). *Single-case research designs: Methods for clinical and applied settings.* Oxford: Oxford University Press.

Kearney, C.A. (1994). Interrater reliability of the Motivation Assessment Scale: Another, closer look. *Journal of the Association for Persons with Severe Handicaps, 19,* 139-142.

Kennedy, C.H., & Haring, T.G. (1993). Combining reward and escape DRO to reduce the problem behavior of students with severe disabilities. *Journal for the Association of Persons with Severe Handicap, 18,* 85-92.

Klier, J., & Harris, S.L. (1977). Self-stimulation and learning in autistic children: Physical or functional incompatibility? *Journal of Applied Behavior Analysis, 10,* 311.

LaGrow, S.J., & Repp, A.C. (1984). Stereotypic responding: A review of intervention research. *American Journal of Mental Deficiency, 88,* 595-609.

Leitenberg, H., Burchard, J.D., Burchard, S.N., Fuller, E.J., & Lysaght, T.V. (1977). Using positive reinforcement to suppress behavior: Some experimental comparisons with sibling conflict. *Behavior Therapy, 8,* 168-182.

Lennox, D.B., & Miltenberger, R.G. (1989). Conducting a functional assessment of problem behavior in applied settings. *Journal of the Association for Persons with Severe Handicaps, 14,* 304-311.

Lennox, D.B., Miltenberger, R.G., & Donnelly, D.R. (1987). Response interruption and DRL for the reduction of rapid eating. *Journal of Applied Behavior Analysis, 20,* 279-284.

Lennox, D.B., Miltenberger, R.G., Spengler, P., & Erfanian, N. (1988). Decelerative treatment practices with persons who have mental retardation: A review of five years of the literature. *American Journal on Mental Retardation, 92,* 492-501.

Levitt, H., & Kaufman, M.E. (1965). Sound-induced drive and stereotyped behavior in mental defectives. *American Journal of Mental Deficiency, 69,* 729-735.

Lovaas, I., Newsom, C., & Hickman, C. (1987). Self-stimulatory behavior and perceptual reinforcement. *Journal of Applied Behavior Analysis, 20,* 45-68.

Lowitz, G.H., & Suib, M.R. (1978). Generalized control of persistent thumb sucking by differential reinforcement of other behaviors. *Journal of Behavior Therapy and Experimental Psychiatry, 9,* 343-346.

Luce, S.C., & Hall, R.V. (1981). Contingent exercise: A procedure used with differential reinforcement to reduce bizarre verbal behavior. *Education and Treatment of Children, 4,* 309-327.

Luiselli, J.K. (1986). Behavior analysis of pharmacological and contingency management interventions for self-injury. *Journal of Behavior Therapy and Experimental Psychiatry, 17,* 275-284.

Luiselli, J.K., Helfen, C.S., Colozzi, G., Donellon, S., & Pemberton, B. (1978). Controlling self-inflicted biting of a retarded child by the differential reinforcement of other behavior. *Psychological Reports, 42,* 435-438.

Mace, F.C., Lalli, J.S., & Lalli, E.P. (1991). Functional analysis and treatment of aberrant behavior. *Research in Developmental Disabilities, 12*, 155-180.

Martin, G., & Pear, J. (1992; 1996). *Behavior modification: What it is and how to do it.* Upper Saddle River, NJ: Prentice-Hall.

Matson, J.L., & Keyes, J.B. (1990). A comparison of movement suppression time-out and DRO with two self-injurious and aggressive mentally retarded adults. *Research in Developmental Disabilities, 11*, 111-120.

Mazaleski, J.L., Iwata, B.A., Vollmer, T.R., Zarcone, J.R., & Smith, R.G. (1993). Analysis of the reinforcement and extinction components in DRO contingencies with self-injury. *Journal of Applied Behavior Analysis, 26*, 143-156.

McClure, J.T., Moss, R.A., McPeters, J.W., & Kirkpatrick, M.A. (1986). Reduction of hand mouthing by a boy with profound mental retardation. *Mental Retardation, 24*, 219-222.

McGonigle, J.J., & Rojahn, J. (1989). An experimental analysis of visual screening and DRO for stereotyped behavior in young children with developmental disabilities. *Journal of the Multihandicapped Person, 2*, 251-270.

McGreevy, P., & Arthur, M. (1987). Effective behavioral treatment of self-biting by a child with Lesch-Nyan syndrome. *Developmental Medicine and Child Neurology, 29*, 536-540.

Miller, B.Y. (1991). *The effect of food on challenging behaviour.* Paper presented at Gregynog, University of Wales conference, May.

Miller, R.S., & Morris, W.N. (1974). The effects of being imitated on children's responses in a marble-dropping task. *Child Development, 45*, 1103-1107.

Mulick, J.A., Leitenberg, H., & Rawson, R.A. (1976). Alternative response training, differential reinforcement of other behavior, and extinction in squirrel monkeys (Samiri sciureus). *Journal of the Experimental Analysis of Behavior, 25*, 311-320.

Newton, J.T., & Sturmey, P. (1991). The Motivation Assessment Scale: inter-rater reliability and internal consistency in a British sample. *Journal of Mental Deficiency Research, 35*, 472-474.

O'Banion, D., Armstrong, B., Cummings, R.A., & Stange, J. (1978). Disruptive behavior: A dietary approach. *Journal of Autism and Developmental Disorders, 8*, 325-337.

O'Brien, S., & Repp, A.C. (1990). Reinforcement-based reductive procedures: A review of 20 years of their use with persons with severe or profound retardation. *Journal of the Association for Persons with Severe Handicap, 15*, 148-159.

Oliver, C., & Head, D. (1993). Self-injurious behaviour: Functional analysis and interventions. In R.S.P. Jones & C.B. Eayrs (Eds.), *Challenging behaviour and intellectual disability: A psychological perspective*. Avon: BILD Publications.

Oliver, C., Murphy, G. H., & Corbett, J. A. (1987). Self-injurious behaviour in people with mental handicap: A total population study. *Journal of Mental Deficiency Research, 31*, 147-162.

Ollendick, T.H., & Matson, J.L. (1978). Overcorrection: An overview. *Behavior Therapy, 9*, 830-842.

Owens, R.G., & Ashcroft, J.B. (1982). Functional analysis in applied psychology. *British Journal of Clinical Psychology, 21*, 181-189.

Pace, G.M., Ivancic, M.T., Edwards, G.L., Iwata, B., & Page, T.J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis, 18*, 249-255.

Patterson, G.R., Jones, R., Whittier, J., & Wright, M.A. (1965). A behaviour modification technique for the hyperactive child. *Behaviour Research and Therapy, 2*, 217-226.

Peterson, R.F., & Peterson, L.R. (1968). The use of positive reinforcement in the control of self-destructive behavior in a retarded boy. *Journal of Experimental Child Psychology, 6*, 351-360.

Poling, A., Miller, K., Nelson, N., & Ryan, C. (1978). Reduction of undesired classroom behavior by systematically reinforcing the absence of such behavior. *Education and Treatment of Children, 1*, 35-41.

Poling, A., & Ryan, C. (1982). Differential reinforcement of other behavior schedules: Therapeutic applications. *Behavior Modification, 6,* 3-21.

Redmon, W.K. (1987). Reduction of physical attacks through differential reinforcement of other behavior. *Journal of Child and Adolescent Psychotherapy, 4,* 107-111.

Reed, J., & Head, D. (1993). The application of functional analysis in the treatment of challenging behaviour. In I. Fleming & B.S. Kroese (Eds.), *People with learning disability and severe challenging behaviour: New developments in services and therapy.* Manchester: Manchester University Press.

Remington, B., & Evans, J. (1988). Basic learning processes in people with profound mental handicaps: Review and relevance. *Mental Handicap Research, 1,* 4-23.

Repp, A.C., Barton, L.E., & Brulle, A.R. (1983). A comparison of two procedures for programming the Differential Reinforcement of Other behaviors. *Journal of Applied Behavior Analysis, 16,* 4, 435-445.

Repp, A.C., & Deitz, D.E.D. (1979). Reinforcement-based reductive procedures: Training and monitoring performance of institutional staff. *Mental Retardation, 17,* 221-226.

Repp, A.C., & Deitz, S.M. (1974). Reducing aggressive and self-injurious behavior of institutionalized retarded children through reinforcement of other behaviors. *Journal of Applied Behavior Analysis, 7,* 313-325.

Repp, A.C., Deitz, S.M., & Deitz, D.E.D. (1976). Reducing inappropriate behaviors in classrooms and in individual sessions through DRO schedules of reinforcement. *Mental Retardation, 14,* 11-15.

Repp, A.C., Deitz, S.M., & Speir, N.C. (1974). Reducing stereotypic responding of retarded persons by the differential reinforcement of other behavior. *American Journal of Mental Deficiency, 79,* 279-284.

Repp, A.C., Felce, D., & Barton, L.E. (1988). Basing the treatment of stereotypic and self-injurious behaviors on hypotheses of their causes. *Journal of Applied Behavior Analysis, 21,* 281-289.

Repp, A.C., Felce, D., & Barton, L.E. (1991). The effects of initial interval size on the efficacy of DRO schedules of reinforcement. *Exceptional Children, 57,* 417-425.

Repp, A.C., Karsh, K.G., Deitz, D.E.D., & Singh, N.N. (1992). A study of the homeostatic level of stereotypy and other motor movements of persons with mental handicaps. *Journal of Intellectual Disability Research, 36,* 61-75.

Repp, A.C., & Singh, N. (1990). *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities.* Sycamore, IL: Sycamore Publishing Co.

Repp, A.C., & Slack, D.J. (1977). Reducing responding of retarded persons by DRO schedules following a history of low-rate responding: A comparison of ascending interval sizes. *The Psychological Record, 3,* 581-588.

Reynolds, G.S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior, 4,* 57-71.

Rolider, A., & Van Houten, R. (1990). The role of reinforcement in reducing inappropriate behavior: Some myths and misconceptions. In A.C. Repp & N. Singh (Eds.), *Perspectives on the use of nonaversive and aversive interventions for persons with developmental disabilities.* Sycamore, IL: Sycamore Publishing Co.

Rollings, J.P., & Baumeister, A.A. (1981). Stimulus control of stereotypic responding: Effects on target and collateral behavior. *American Journal of Mental Deficiency, 86,* 67-77.

Rubin, G., Griswald, K., Smith, I., & DeLeonarda, C. (1972). A case study in the remediation of severe self-destructive behavior in a 6-year-old mentally retarded girl. *Journal of Clinical Psychology, 28,* 424-426.

Saloviita, T. (1988). Elimination of self-injurious behaviour by brief physical restraint and DRA. *Scandinavian Journal of Behaviour Therapy, 17,* 55-63.

Schuler, A.L., & Goetz, C. (1981). The assessment of severe language disabilities: Communicative and cognitive considerations. *Analysis and Intervention in Developmental Disabilities, 1,* 333-346.

Singh, N.N., Dawson, M., & Manning, P. (1981). Effects of spaced responding DRL on the stereotyped behavior of profoundly retarded persons. *Journal of Applied Behavior Analysis, 14,* 521-526.

Singh, N.N., & Repp, A.C. (1988). The behavioural and pharmacological management of problem behaviours in people with mental retardation. *The Irish Journal of Psychology, 9,* 264-285

Sisson, L.A., Van Hasselt, V.B., Hersen, M., & Aurand, J.C. (1988). Tripartite behavioral intervention to reduce stereotypic and disruptive behaviors in young multihandicapped children. *Behavior Therapy, 19,* 503-526.

Skinner, B.F. (1938). *The behavior of organisms: An experimental analysis.* Englewood Cliffs, NJ: Prentice-Hall.

Skinner, B.F. (1953). *Science and human behavior.* New York: Macmillan.

Skinner, B.F. (1948). "Superstition" in the pigeon. *Journal of Experimental Psychology, 38,* 168-172.

Smith, M.D. (1987). Treatment of pica in an adult disabled by autism by differential reinforcement of incompatible behavior. *Journal of Behavior Therapy and Experimental Psychiatry, 18,* 285-288.

Steege, M., Wacker, D.P., Berg, W.K., Cigrand, K.K., & Cooper, L.J. (1989). The use of behavioral assessment to prescribe and evaluate treatments for severely handicapped children. *Journal of Applied Behavior Analysis, 22,* 23-33.

Sturmey, P. (1996). *Functional analysis in clinical psychology.* Chichester: John Wiley & Sons.

Sulzer-Azaroff, B., & Mayer, G.R. (1977). *Applying behavior-analysis procedures with children and youth.* New York: Holt, Rinehart and Winston.

Tarpley, H.D., & Schroeder, S.R. (1979). Comparison of DRO and DRI on rate of suppression of self-injurious behavior. *American Journal of Mental Deficiency, 84,* 188-194.

Thelen, E. (1980). Determinants of amounts of stereotyped behavior in normal human infants. *Ethology and Sociobiology, 1,* 141-150.

Tierney, I.R., McGuire, R.J., & Walton, H.J. (1979). Reduction of stereotyped body-rocking using variable time reinforcement: Practical and theoretical implications. *Journal of Mental Deficiency Research, 23,* 175-185.

Toogood, A. (1996). *A comparison of functional analysis methods.* Unpublished doctoral dissertation, University of Wales, Bangor.

Topping, J.S., Graves, A.J., & Moss, J.D. (1975). Response elimination in elementary and special education school children. *Psychological Record, 25,* 567-572.

Uhl, C.N., & Garcia, E.E. (1969). Comparison of omission with extinction in response elimination in rats. *Journal of Comparative and Physiological Psychology, 69,* 554-562.

Uhl, C.N., & Sherman, W.O. (1971). Comparison of combinations of omission, punishment, and extinction methods in response elimination in rats. *Journal of Comparative and Physiological Psychology, 74,* 59-65.

Van Houten, R., & Rolider, A. (1991). Applied behavior analysis. In J.L. Matson & J.A. Mulick (Eds.), *Handbook of mental retardation,* 2nd edition. New York: Pergamon.

Vaughan, M.E., & Michael, J.L. (1982). Automatic reinforcement: An important but ignored concept. *Behaviorism, 10,* 217-228.

Vollmer, T.R. (1994). The concept of automatic reinforcement: Implications for behavioral research in developmental disabilities. *Research in Developmental Disabilities, 15,* 187-207.

Vollmer, T.R., & Iwata, B.A. (1992). Differential reinforcement as treatment for behavior disorders: Procedural and functional variations. *Research in Developmental Disabilities, 13,* 393-417.

Vollmer, T.R., & Iwata, B.A. (1991). Establishing operations and reinforcement effects. *Journal of Applied Behavior Analysis, 24,* 279-291.

Vollmer, T.R., Iwata, B.A., Zarcone, J.R., Smith, R.G., & Mazaleski, J.L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis, 26,* 9-21.

Wagaman, J.R., Miltenberger, R.G., & Williams, D.E. (1995). Treatment of a vocal tic by differential reinforcement. *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 35-39.

Wahler, R.G., & Fox, J.J. (1981). Setting events in applied behavior analysis: Towards a conceptual and methodological expansion. *Journal of Applied Behavior Analysis, 14,* 327-338.

Warren, S.A., & Burns, N.R. (1970). Crib confinement as a factor in repetitive and stereotyped behavior in retardates. *Mental Retardation, 8,* 25-29.

Whitaker, S. (1992). The reduction of low frequency aggression in a woman with learning difficulties. *Behavioural Psychotherapy, 20,* 373-379.

Wieseler, N.A., Hanson, R.H., Chamberlain, T.P., & Thompson, T. (1985). Functional taxonomy of stereotypic and self-injurious behavior. *Mental Retardation, 25,* 230-234.

Wolery, M., Kirk, K., & Gast, D.L. (1985). Stereotypic behavior as a reinforcer: Effects and side effects. *Journal of Autism and Developmental Disorders, 15,* 149-161.

Wolfensberger, W., & Thomas, S. (1982). *The principle of normalisation in human services.* Toronto: National Institute on Mental Retardation.

Woods, T.S. (1983). The selective suppression of a stereotypy in an autistic child: A stimulus control approach. *Behavioural Psychotherapy, 11,* 235-248.

Young, J.A., & Wincze, J.P. (1974). The effects of the reinforcement of compatible and incompatible alternative behaviors on the self-injurious and related behaviors of a profoundly retarded female adult. *Behavior Therapy, 5,* 614-623.

Zarcone, J.R., Rodgers, T.A., & Iwata, B.A. (1991). Reliability analysis of the Motivation Assessment Scale: A failure to replicate. *Research in Developmental Disabilities, 12,* 349-360.

Zeiler, M.D. (1970). Other behavior: Consequences of reinforcing not responding. *The Journal of Psychology, 74,* 149-155.

Zeiler, M.D. (1979). Reinforcing the absence of fixed-ratio performance. *Journal of the Experimental Analysis of Behavior, 31,* 321-332.

Appendix A

Behaviour recording sheets

- Momentary time-sampling sheet (Experiments 1, 2, 5, 6, 7, 8)
- Duration recording sheet (Experiment 3)
- Duration recording sheet (Experiment 4)

10 SECOND MOMENTARY TIME SAMPLING SHEET
(EXPERIMENTS 1, 2, 5, 6, 7, 8)

Date............................   Client's Name...............................

Please put a tick √ every time you see any of the following behaviours:
Please put an x if the behaviour is not present.
Please put a zero (o) if you can't see the person at the time of the signal

Behaviour 1 (specify)

....

....

Time recording began...............................

1	2	3	4	5
☐	☐	☐	☐	☐
6	7	8	9	10
☐	☐	☐	☐	☐
11	12	13	14	15
☐	☐	☐	☐	☐
16	17	18	19	20
☐	☐	☐	☐	☐
21	22	23	24	25
☐	☐	☐	☐	☐
26	27	28	29	30
☐	☐	☐	☐	☐

Inter rater measure for this session?............

Behaviour 2 (specify)

....

....

Time recording began...............................

1	2	3	4	5
☐	☐	☐	☐	☐
6	7	8	9	10
☐	☐	☐	☐	☐
11	12	13	14	15
☐	☐	☐	☐	☐
16	17	18	19	20
☐	☐	☐	☐	☐
21	22	23	24	25
☐	☐	☐	☐	☐
26	27	28	29	30
☐	☐	☐	☐	☐

Inter rater measure for this session?............

Behaviour 1 (specify)

....

....

Time recording began...............................

1	2	3	4	5
☐	☐	☐	☐	☐
6	7	8	9	10
☐	☐	☐	☐	☐
11	12	13	14	15
☐	☐	☐	☐	☐
16	17	18	19	20
☐	☐	☐	☐	☐
21	22	23	24	25
☐	☐	☐	☐	☐
26	27	28	29	30
☐	☐	☐	☐	☐

Inter rater measure for this session?............

Behaviour 2 (specify)

....

....

Time recording began...............................

1	2	3	4	5
☐	☐	☐	☐	☐
6	7	8	9	10
☐	☐	☐	☐	☐
11	12	13	14	15
☐	☐	☐	☐	☐
16	17	18	19	20
☐	☐	☐	☐	☐
21	22	23	24	25
☐	☐	☐	☐	☐
26	27	28	29	30
☐	☐	☐	☐	☐

Inter rater measure for this session?............

SHEET USED TO RECORD DURATIONS OF .

LOWER-RATE BEHAVIOURS (EXPERIMENT 3)

Session:

Start time:

Target behaviour:

Time from start of sessions (in seconds)	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59
0						
1						
2						
3						
4						
0						
1						
2						
3						
4						
0						
1						
2						
3						
4						
0						
1						
2 -						
3						
4						

SHEET USED TO RECORD DURATIONS OF
LOWER-RATE BEHAVIOURS (EXPERIMENT 4)

Session:

Start time:

Target behaviour:
Monitored behaviour:

Time from start of sessions (in seconds)	Target behaviour						Monitored behaviour						On-task behaviour					
	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59
0																		
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		
18																		
19																		
20																		
21																		
22																		
23																		
24																		
25																		
26																		
27																		
28																		
29																		
30																		

Appendix B

Rationale for the approximations of the randomization test

APPROXIMATIONS OF THE RANDOMIZATION TEST

"For single-case research, t and F can be readily used with the proviso that randomization of conditions to occasions must be met." (Kazdin, 1982, p. 328)

The randomization test (Edgington, 1969) is suitable for use in single-case designs where conditions are frequently alternated. The test aims to determine whether behaviour under one condition differs significantly from behaviour under the other condition or conditions. The null hypothesis states that any difference in behaviour between conditions is due to chance; the randomization test calculates the probability of obtaining a difference between treatments as large as the one obtained. It is essential that the various conditions are randomly assigned, and that the total number of sessions is determined prior to experimentation.

Kazdin (1982) notes that although this test is useful when there are a small number of data points, in cases of more than ten sessions the calculations are so large and complex as to be unmanageable. However, t tests (for two conditions) and F tests (for more than two conditions) can be used as approximations to the randomization test. Ordinarily, the use of t and F in single-case designs is not advised, as single-case data tend to show serial dependency. That is, data on one day often predict the data on the next day and so on: data are correlated and are not independent. As independent data is an assumption of t and F tests, in cases where there is serial dependency these tests cannot be used. However, the use of t and F as approximations to the randomization test is not effected by dependency, because in the alternating treatments design conditions are randomly presented. Therefore these approximations can be used in such cases, regardless of the presence of serial dependency. The alternating treatments design lends itself very well to analysis by these approximations, as it consists of randomly and rapidly alternated conditions.

Edgington, E.S. (1969). *Statistical inference: The distribution-free approach.* New York: McGraw Hill.

Kazdin, A.E. (1982). *Single-case research designs: Methods for clinical and applied settings.* Oxford: Oxford University Press.

Appendix C

Questionnaire given to subjects at the end of Experiment 7

<u>Debriefing</u>

Thank you for participating in this experiment. The study is investigating the effect of reinforcers (in this case money) on an individual's ability to work out which segment to avoid.

A full debriefing will be sent to your pigeon-hole once the whole experiment has been completed. Could you please indicate on the diagram below which area you thought you were to avoid in order to gain points.

Appendix D

Materials used in Experiment 9

- Instructions
- Questionnaire
- Distraction script

INSTRUCTIONS FOR SUBJECTS IN WHOLE-INTERVAL DRO CONDITION

Using reinforcers to reduce challenging behaviour: A simulation
Experiment run by Beth Miller

Summary of experiment

In this experiment you are asked to simulate the implementation of a clinical technique by 'rewarding' the actions of an actor on a video-tape. 'Rewards' are given by pressing a computer key.

You are asked to reward the 'client' - Danny - for every <u>**35 seconds**</u> that he does *not* rub his hand. This length of time has been chosen by using the average length of time he habitually refrains from this behaviour.

When you get to this point, please ask me to show you a clip of Danny on the video, to give you an idea of what his hand-rubbing looks like.

To monitor the length of time that Danny avoids hand-rubbing, you will use a stop-watch on the computer. You will start timing when Danny stops hand-rubbing, and watch him continually.

If he doesn't rub his hand for the whole 35 seconds, then you should stop the watch, press the 'Reward' button, and reset the watch to zero. You should then carry on with the procedure for another 35 seconds.

If you start the watch but Danny rubs his hand before the 35 seconds are up, you should stop the watch, reset it back to zero, and wait for him to cease his hand-rubbing. Once he does stop, you can start timing again.

You should never be timing Danny if he is rubbing his hand.

First of all you will be given a practice session which will go on for five minutes. This will be followed by an experimental session which will go on for fifteen minutes.
When stopping the watch, rewarding, resetting and restarting the watch, you should try and be as quick as possible.

Please note: If any of the instructions are not clear, or if you have any questions, please ask me.

SUMMARY OF INSTRUCTIONS FOR WHOLE-INTERVAL DRO

- Start the stop-watch if Danny isn't rubbing his hands.

- Watch Danny continuously while the video is playing.

- If Danny goes for 35 seconds <u>without hand-rubbing</u>:
  - Stop the watch
  - Press the 'Reward' button (space-bar)
  - Reset the watch to zero
  - If he's still *not* rubbing his hands, start timing and watching again
  - If he *is* rubbing his hands, wait till he stops, then start timing again.

- If Danny <u>rubs his hand</u> before the 35 seconds are up
  - Stop the watch
  - Reset the watch to zero
  - Wait for him to cease his hand-rubbing
  - Once he does stop, you can start timing again

<u>NB</u> You should never be timing Danny if he is rubbing his hand.

INSTRUCTIONS FOR SUBJECTS IN MOMENTARY DRO CONDITION

Using reinforcers to reduce challenging behaviour: A simulation
Experiment run by Beth Miller

Summary of experiment
In this experiment you are asked to simulate the implementation of a clinical technique by 'rewarding' the actions of an actor on a video-tape. 'Rewards' are given by pressing a computer key.

You are asked to use the stop-watch to time sessions of __35 seconds__. This length of time has been chosen by using the average length of time the 'client' - Danny - habitually refrains from rubbing his hand.

When you get to this point, please ask me to show you a clip of Danny on the video, to give you an idea of what his hand-rubbing looks like.

This means that you start the watch at the start of the video. When it reaches 35 seconds, you should stop it, and look at the same moment to see what Danny is doing.

If at the moment the 35 seconds were up he was NOT rubbing his hand, you should press the 'Reward' button, then reset the watch to zero. You should then start timing another 35 seconds.

If at the moment the 35 seconds were up he was rubbing his hand, you should stop the watch, reset it to zero, and time another 35 seconds.

First of all you will be given a practice session which will go on for five minutes. This will be followed by an experimental session which will go on for fifteen minutes. When stopping the watch, rewarding, resetting and restarting the watch, you should try and be as quick as possible.

Please note: If any of the instructions are not clear, or if you have any questions, please ask me.

SUMMARY OF INSTRUCTIONS FOR MOMENTARY DRO

- Start the stop-watch and time for 35 seconds.

- At the end of the 35 seconds
    - Stop the watch
    - At the same moment, look to see what Danny is doing on the video.

- If Danny <u>was not</u> rubbing his hand *at that moment* you looked at him:
    - Press the 'Reward' button (Spacebar)
    - Reset the watch to zero
    - Start timing another 35 seconds

- If Danny <u>was</u> rubbing his hand *at that moment* you looked at him:
    - Reset the watch to zero
    - Start timing another 35 seconds

QUESTIONNAIRE GIVEN TO SUBJECTS AT END OF EXPERIMENT

Subject number
Session

Question for the end of the experiment

Many thanks for taking part in this study.

Please state how difficult you found it to carry out the instructions you were given, by ticking in the appropriate box, below:

Very easy ☐   Easy ☐   Quite easy ☐

Quite difficult ☐   Difficult ☐   Very difficult ☐

Please can you give some more detailed information about your thoughts on what you were asked to do in this experiment?

Many thanks for participating in this study. I will send you full details of the study in writing when all experimental subjects have been run.

DISTRACTION SCRIPT

Beth will give instructions to subjects. Subjects have a 5 minute practice.

Beth will ring Becca saying.....

" I'm starting with my subject now...Can you carry on with that self-assessment stuff? Thanks"

Beth puts phone through to Becca. Becca will start timing immediately after putting the phone down.

Distraction 1

At 3 minutes, Becca rings Beth....

Beth: *"hello"*

Becca: *"Oh hello Beth it's Becca, have you got a minute?"*

Beth: *"Yes..."*

Becca: *"You know that QuickMail you sent yesterday?"*

Beth: *"Yes"*

Becca: *" Well I've made the changes you suggested..."*

Beth: *"Right"*

Becca: *"And I'm waiting to hear from Alun on one thing..."*

Beth: *"OK..."*

Becca: *"But I didn't understand the last bit about weaknesses, can you explain?"*

Beth: *"Oh yes, I meant we haven't really taken on board that bit that's mentioned in section 2 of the HEFCW guidelines"*

Becca: *"Is that the big handout with the jargon in?"*

Beth: *"Yes, have you got it there?"*

Becca: *"Well I've got a copy here, 1994-1995, is that it?*

Beth: *"No, it's the 1995-96 ones"*

Becca: *"Oh right - have I definitely got that then?"*

Beth: *"Maybe not. David Roberts sent it out in June"*

Becca: *"It might be filed away, can you give me a minute?   _*

Beth: *"OK, I'll hang on while you look for it.*

(PAUSE FOR TEN SECONDS)

Becca: *"OK! Found it!"*

Beth: *"Good, right well you see on page 13, down the bottom, it says 'judgements about the effectiveness of teaching and learning activities might be presented best in terms of perceived strength and weaknesses'."*

Becca: *"Yep... got it"*

Beth: *"That's what I meant, I was worried we didn't really say anything about our perceived weaknesses"*

Becca: *"We just go on about how great we are all the time"*

Beth: *"Yes, that's right, loads of stuff about our strengths but I'm worried it might look a bit odd if we don't say that we know where our weaknesses are. Have any of the drafts that you've had back mentioned any weaknesses at all that you can see?"*

Becca: *"I haven't gone through them yet - do you want to go through them with me now?"*

Beth: *"No, I can't, I'm running the experiment now. Can't you have a look through them and get back to me?"*

Becca: *"Well the problem is there's some stuff we've got to finalise today..."*

Beth: *"I know, we do need it to be finished quickly but I'm running subjects here and got millions more to do"*

Becca: *"It's only a little thing and I've more or less finished it"*

Beth: *"Ok, well can you send me a QuickMail about it - don't phone"*

Becca: *"It'd be much quicker if I just brought it in to show you..."*

Beth: *"Ok, if it has to go today I'll quickly check it now, is it brief?"*

Becca: *"Yes"*

Beth: *" Ok then, see you in a minute."*

Becca hangs up and starts timing again. After one minute as passed, she knocks on Beth's door:

Distraction 2

Beth : *"Come in..."*.

Becca enters.

Beth says to subjects: *" ...sorry about this...could you please carry on with what you're doing..."*

Beth & Becca stand behind the subject and have the following conversation

Becca : *"Hi...sorry..."* , while handing Beth piece of paper (pen profile) to check.

 Beth : *"OK...right"* as she reads over paper.

Beth: *"Has Fergus seen this?"*

Becca: *"No, I didn't know I should show it to him..."*

Beth (bit cross) *"Yes! He has to see everything that we're going to send to the assessors. What about the other profiles, did he not see those?"*

Becca: *"Well I passed them on to Alun after you'd seen them and he gave them back, didn't say anything about Fergus having to see them... Does it matter? We've all seen them? Would it make a big difference?"*

Beth: *"I don't know, I wanted everything to be checked 100 times before we sent it out, I just want to make sure that everything is completely right."*

Becca: *"Well what about all the supporting stuff that I'm collecting now?"*

Beth: *"Oh yes, anything that won't already have been checked should go out to me and Alun and Fergus, what stuff have you got so far?"*

Becca: *"Can't remember all of it, I can have a look if you like, there's some lecture handouts, some lists of phone numbers, the syllabuses for most of the courses, project guidelines, that sort of thing..."*

Beth: *"Well would you be able to do three copies of everything and send them out to us? When were you wanting to send them out by?"*

Becca: *"Oh not till everything else is ready to go out, it won't go out separately will it?"*

Beth: *"No, spose not, well in that case there's no rush, just as supporting documents come in you could copy them and send them out to us, then we can check everything, make sure there's nothing going out that looks a bit dodgy..."*

Becca: *"I thought you wanted us to show off our weaknesses..."*

Beth: *"Oh ha ha Becca, yes very funny..."*

Becca: *"There's one more profile I'm chasing up - I'm going to see if I can get it now, if I do can you check it before then end of today?"*

Beth: *"Yes ok"*

Becca: *"Ok, see you later Beth. Bye."*

Beth: *"Right. Bye".*

Becca leaves.

Beth says to subjects: *" ...sorry about that..."*

Becca starts timing again. Three minutes later, she rings Beth.

Distraction 3

Beth: *"Hello?"*

Becca: *"Hello, I'm phoning about the research methods course this year...."*

Beth: *"well I really can't talk about this now...."*

Becca: *"It won't take long, it's a really quick query..."*

Beth: *"I'm actually in the middle of an experiment.!"*

Becca: *"Well how was I meant to know?"*

Beth: *"I asked not to be interrupted!! They shouldn't be putting any calls through to me at all really. "*

Becca: *"Well I've got a bit of a query about the syllabus..."*

Beth: *"You'll just have to sort it out yourself!! "*

Becca: *"But I was told you were the person to ask...."*

Beth: *"I'm sorry, I really don't think it's got anything to do with me. "*

Becca: *"Well I asked Alun and he said you'd know.."*

Beth: *"I don't know why he said that. There's no way I should be involved in dealing with that.*

Becca: *"Not even queries about assessment?"*

Beth: *"No, it isn't up to me at all and I just can't quite work out why he's trying to bring me in on it anyway. "*

Becca: *"So can you tell me who I can ask then?"*

Beth: *"No. Look, I can't talk about this now. As I said I'm in the middle of an experiment."*

Becca: *"Bye"*

Beth: *"Bye"*

Beth puts phone down.

Three minutes later Becca returns, again knocking on the door before entering

Distraction 4

Beth : *"Come in...".*

Becca ,who is carrying another profile, says ....*"Here's that last profile I said..."*

Beth : *"Can you just leave it there Becca and I'll look at it later".*

Becca: *"Ok. I just saw Alun and he says can you give him a ring when you get a minute?"*

Beth: *"Ok, nearly finished here"*

Becca goes out

Beth (to subject): *"Sorry about that."*

Appendix E

ANOVA Summary Tables

CHAPTER 3 TABLES

Table 1   Analysis of variance of differences in percentage of target behaviour between whole-interval DRO, momentary DRO and momentary DRO/2 (randomization test) in Experiment 5.

Within-subjects factor: Type of schedule.

Source	SS	MS	DF	F	p value
Schedule	1178.74	589.37	2	0.99	0.3819
Error	20844.039	595.544	35		

Table 2   Analysis of variance of differences in percentage of monitored behaviour between whole-interval DRO, momentary DRO and momentary DRO/2 (randomization test) in Experiment 5.

Within-subjects factor: Type of schedule.

Source	SS	MS	DF	F	p value
Schedule	1520.969	760.485	2	2.839	0.072
Error	9375.49	267.871	35		

Table 3   Analysis of variance of differences in percentage of target behaviour between whole-interval DRO, momentary DRO and momentary DRO/2 (randomization test) in Experiment 6.

Within-subjects factor: Type of schedule.

Source	SS	MS	DF	F	p value
Schedule	2621.067	1310.533	2	12.291	0.0002
Error	2878.8	106.622	27		

Table 4   Analysis of variance of differences in percentage of monitored behaviour between whole-interval DRO, momentary DRO and momentary DRO/2 (randomization test) in Experiment 6.

Within-subjects factor: Type of schedule.

Source	SS	MS	DF	F	p value
Schedule	182.893	91.446	2	3.774	0.0359
Error	654.176	24.229	27		

CHAPTER 4 TABLES

Table 5   Analysis of variance of differences in amount of target behaviour across schedules and phases in Experiment 7.

Between-subjects factor: DRO schedule.

Within-subjects factor: Phase of experiment.

Source	SS	MS	DF	F	p value
Schedule	83.37	83.37	1	0.56	0.461
Error	4185.61	149.49	28		
Phase	1341.10	447.03	3	6.47	0.001
Schedule x Phase	401.81	133.94	3	1.94	0.130
Error	5806.77	69.13	84		

Table 6   Analysis of variance to determine which of the phase comparisons in whole-interval DRO were responsible for the significant difference shown by the mixed ANOVA, above (Experiment 7).

Within-subjects factor: Phase of experiment.

Source	SS	MS	DF	F	p value
Phase	208.232	69.411	3	0.67	0.5738
Error	5798.788	103.55	56		

Table 7   Analysis of variance to determine which of the phase comparisons in momentary DRO were responsible for the significant difference shown by the mixed ANOVA, above (Experiment 7).

Within-subjects factor: Phase of experiment.

Source	SS	MS	DF	F	p value
Phase	1534.679	511.56	3	6.831	0.0005
Error	4193.593	74.886	56		

Table 8 - Analysis of variance of differences across schedule and phase, with regard to whether subjects had correctly identified the designated area in Experiment 7. Between-subjects factors: Schedule and identification of designated area. Within-subjects factor: Phase of intervention (Baseline phases were omitted from this analysis)

Source	SS	MS	DF	F	p value
Schedule	85.91	85.91	1	0.59	0.454
Identification	1007.15	1007.15	1	6.87	0.014
Schedule x Identification	1.20	1.20	1	0.01	0.929
Error	3812.54	146.64	26		
Phase	279.89	139.95	2	2.58	0.085
Phase x Schedule	202.12	101.06	2	1.86	0.165
Phase x Identification	325.29	162.65	2	3.00	0.058
Phase x Schedule x Identification	27.03	13.51	2	0.25	0.780
Error	2817.88	54.19	52		

CHAPTER 6 TABLES

Table 9   Analysis of variance of subjects' difficulty of task scores in Experiment 9. Between-subjects factors: DRO schedule and condition.

Source	SS	MS	DF	F	p value
Schedule	0.01667	0.01667	1	0.01852	0.8922
Condition	0.15	0.15	1	0.16667	0.6846
Schedule x Condition	0.01667	0.01667	1	0.01852	0.8922
Error	50.4	0.9	56		

Table 10   Analysis of variance of the number of rewards subjects gave in Experiment 9. Between-subjects factors: DRO schedule and condition.

Source	SS	MS	DF	F	p value
Schedule	290.4	290.4	1	960.37795	0.0001
Condition *	0	0	1	0	•
Schedule x Condition	0.26667	0.26667	1	0.88189	0.3517
Error	16.93333	0.30238	56		

*   The mean number of rewards given in silent and distracting conditions were identical (14.2). This results in the zero values seen in this row.

Table 11 Analysis of variance of percentage error scores of number of rewards in Experiment 9.
Between-subjects factors: DRO schedule and condition.

Source	SS	MS	DF	F	p value
Schedule	115.759	115.759	1	12.218	0.0009
Condition	23.438	23.438	1	2.474	0.1214
Schedule x Condition	23.437	23.437	1	2.474	0.1214
Error	530.568	9.474	56		

Table 12 Analysis of variance of percentage of accuracy of timing scores in Experiment 9. Between-subjects factors: DRO schedule and condition.

Source	SS	MS	DF	F	p value
Schedule	18302.028	18302.028	1	19.535	0.0001
Condition	159.789	159.789	1	0.171	0.6812
Schedule x Condition	218.829	218.829	1	0.234	0.6308
Error	52465.927	936.892	56		