The Behavior Analyst Certification Board (BACB) Guidelines for Responsible Conduct for Behavior Analysts (BACB, 2001) state clearly that behavior analysts should conduct functional assessment as a part of their work with a client. The purpose of the present paper is to explore the ethical implications of conducting or not conducting a functional assessment. The discussion will focus on the case of self-injurious behavior in people with developmental disabilities. Self-injury is a major clinical problem that can result in significant injury and even death, and has been associated with negative outcomes for carers such as stress (Hastings, 2002). Epidemiological surveys suggest that up to 20% of adults with mental retardation engage in some form of self-injurious behavior (Collacott, Cooper, Branford, & McGrother, 1998; Emerson et al., 2001; Oliver, Murphy, & Corbett, 1987). Thus, self-injury is a problem of considerable applied significance falling within the purview of Applied Behavior Analysis (ABA; Baer, Wolf, & Risley, 1968).

There are a number of principles applicable to the question of whether treating self-injury with an intervention based on a pre-treatment functional assessment is ethically superior to implementing an intervention without a functional assessment. Those that will be discussed here are: the right to effective treatment, the application of least restrictive treatment alternatives, social validity, and the importance of a constructional approach within ABA. Before discussing these ethical issues, it is important to define some terminology to be used in this paper. ‘Functional Assessment’ will be used to refer to a whole variety of methods that might be used to inform hypotheses about the function of self-injury, whereas ‘Analog Assessment’ will be reserved specifically to refer to experimental functional analysis (cf. Iwata, Dorsey, Slifer, Bauman, & Richman, 1982). Some of the ethical issues relating to experimental and non-experimental methods of functional assessment are different and so this is an important distinction to maintain.

The contrast in the discussion below is between treatment based on functional assessment, and an approach referred to as ‘Behavior Modification’. Behavior Modification is defined here as the application of behavior change technologies (i.e., via reinforcement...
or punishment techniques) without reference to the analytic dimension of ABA (Baer et al., 1968). One of the defining features of ABA is that it is analytic in nature: involving a reliable demonstration of the events responsible for the occurrence and non-occurrence of behavior (Baer et al.). This central distinction of whether treatments are based on hypotheses about their causes (Repp, Felce, & Barton, 1988) has ethical implications for the treatment of self-injury in individuals with developmental disabilities.

Effective Treatment for Self-Injury

Effectiveness of treatment can be defined in a number of ways. At the basic level, it is certainly possible to identify reports in the research literature demonstrating that behavior modification procedures for self-injury can lead to clinically salient outcomes (see Repp, Singh, Olinger, & Olson, 1990, for a review). Similarly, there are reports demonstrating that clinically salient outcomes are achieved by carrying out a functional assessment and deriving a treatment for self-injury from hypotheses about the processes maintaining the behavior (e.g., Carr & Durand, 1985). The central ethical point is that clients should be offered the most effective treatments available.

Probably the most convincing evidence that addresses the question of relative effectiveness comes from research employing meta-analytic methods to explore variables that moderate the effectiveness of treatment for severe behavior problems such as self-injury. In reviewing a large number of single case experiments reporting treatment of severe behavior problems, several authors have used a standard statistic called the Percentage of Non-Overlap ping Data points (PND) (Scruggs & Mastropieri, 1998) to quantify treatment effects within each experiment. A very effective treatment will have a PND of 1.0 (i.e., there is no overlap between data points during baseline and treatment phases). Further characteristics of each experiment can be coded (e.g., age of client, severity of disability, whether a pre-treatment functional assessment was conducted) and then explored as moderators of the size of treatment effects as measured using the PND. A consistent finding across meta-analytic reviews is that functional assessment significantly increases the effectiveness of the behavioral treatment of severe behavior problems (Carr et al., 1999; Didden, Duker, & Korzilius, 1997; Scotti, Evans, Meyer, & Walker, 1991; Sternberg, Taylor, & Babkie, 1994).

A difficulty with many studies of self-injury is that they do not report data on generalization and maintenance of treatment effects (Scotti, Evans, Meyer, & Walker, 1991). It is important to ask about the effectiveness of treatment both in general and in terms of their success according to these two dimensions. An effective treatment must generalize to new situations. This is something that can be made more likely by attending to generalization issues (e.g., implementing treatment in multiple contexts). However, planning for generalization could be applied similarly to a treatment based on functional assessment or one based on behavior modification. Perhaps a more crucial issue is actually the long-term maintenance of treatment effects. Studies of long-term maintenance of interventions for self-injury are very rare. A notable exception is Durand and Carr’s (1992) research, which showed that Functional Communication Training (FCT) and timeout were initially successful to a similar extent, but that the effects of FCT were more likely to maintain. In terms of generalization, there were also data suggesting that FCT but not timeout also generalized to teachers naïve about the interventions.

Without a large evidence base, we need to return to an understanding of a key principle of behavior to explore whether we would expect behavior modification or a treatment based on functional assessment to be more effective in the longer term. A key criticism of early behavior modification was the potential for the problem of response substitution (Kazdin, 1982). As behavior modification strategies do not address the ‘underlying causes’ of behavior problems, then simply reducing the behavior may have unpredictable effects with apparently new behavior problems emerging to take their place. This potential problem can be understood with reference to the construct of the response class.

A response class consists of behaviors that serve the same function (i.e., are functionally but not necessarily topographically equiva-
lent). Treatments for self-injury such as FCT (Bird, Dores, Moniz, & Robinson, 1989; Carr & Durand, 1985) make direct use of this concept of functionally equivalent responses. In FCT, the social environment is arranged such that self-injury is not reinforced but a functionally equivalent socially appropriate behavior (e.g., vocal or signed request for attention or help with difficult tasks) is reinforced. The idea is that as the more ‘acceptable’ member of the response class more reliably leads to reinforcement, self-injury is less likely to occur (Scotti, Evans, Meyer, & DiBenedetto, 1991). Principles related to the matching law (McDowell, 1988; Myerson & Hale, 1984; Pierce & Epling, 1995) would predict such a state of affairs, as organisms will tend to behave in a manner that maximises reinforcement.

Returning to the question of the relative effectiveness of a functional approach versus behavior modification, one would predict that treatment effects achieved through behavior modification would be less likely to maintain. This would be for two main reasons. First, there would be no consistent selection of an appropriate alternative response, and so the emergence of behavior from the same response class to achieve the same function would be unpredictable. The ‘new’ behavior in these circumstances could well be another self-injury topography or another problem behavior. Second, without an explicitly developed alternative response that is functional in a range of environments (note that this is also a generalization issue), a client would be very likely to revert to self-injury once the behavioral programme in place was ended or once its effective implementation began to deteriorate.

Treatment fidelity, or the extent to which a procedure is implemented accurately, is another level at which we might consider the effectiveness of treatments for self-injury. A final dimension is the likely effect of beginning treatment on frequency of self-injury due to the phenomenon of the extinction burst. When reinforcement is removed, the affected behavior tends to occur at higher frequencies for a brief period. The significance of both treatment fidelity and the phenomenon of the extinction burst in the present discussion can be illustrated by describing a research study conducted by McConnachie and Carr (1997). McConnachie and Carr (1997) found that interventions for self-injury based on a FCT approach led to a reduced extinction burst when compared to escape extinction. Given that one would wish ethically to minimize the occurrence of self-injury, an intervention based on a functional assessment may be more effective at this level. A further key result of the study was that when staff were not being monitored closely by supervisors, those using escape extinction were less consistent in applying their intervention. McConnachie and Carr suggest that these two phenomena are linked in that staff intervention behavior is maintained by avoidance contingencies. Challenging behaviours such as self-injury are aversive (Mosman, Hastings, & Brown, 2002) and through avoiding them, counter-habilitative staff behaviors can be negatively reinforced (Hastings & Remington, 1994).

**Least Restrictive Treatment Procedures**

A second ethical principle guiding the work of behavior analysts is that the least restrictive treatment alternative should be used (Johnston & Sherman, 1993). The use of punishment procedures is not explicitly ruled out, although behavior analysts are encouraged to use reinforcement procedures whenever possible. The underlying goal of a behavior modification approach to self-injury is to remove or reduce the frequency of the behavior. This approach might also be called eliminative (Goldiamond, 1974). An approach based on a functional assessment is constructional (Goldiamond), as new behaviours are learned to replace self-injury. Given the focus on learning of new behavior, or development of behaviors already in the client’s repertoire, treatment based on a functional assessment is by definition likely to be less restrictive and thus ethically preferable.

There may be an argument, however, for the use of restrictive behavior modification techniques when self-injury is extreme and/or immediately life-threatening. Punishment correctly applied can be effective in rapidly reducing self-injury. For example, Romanetsy and Goren (1975) showed that contingent electric shock could be used to suppress self-injury. Thus, there may be situations in which one could argue that the first goal of treat-
ment is to reduce the risk of self-injury and thus it is ethical to use a restrictive procedure.

Social Validity

A third ethical issue of relevance to the treatment of self-injury is selecting treatments that lead to socially valued outcomes via socially valued means. The BACB guidelines are not particularly clear on this issue, referring in the section on research issues to the expectation that behavior analysts act in a manner that maintains the dignity of participants. Socially valued outcomes are also explicit in the description of ABA as Applied (Baer et al., 1968). The concept of the social validity of treatments has, however, also been an important issue in behavior analysis for some years.

Behavior analysts have espoused the principle that data on acceptability of interventions (to staff, families, wider society, and people with developmental disabilities themselves) should help to guide treatment selection. This criterion has been given the general label of ‘social validity’, or the views of consumers on the importance, acceptability and effectiveness of interventions (Wolf, 1978). Probably the most extensively studied dimension of social validity is the social significance of behavioral intervention procedures themselves (Wolf), operationalized as the appropriateness, fairness, and reasonableness of procedures applied to ameliorate a behavioral problem (Kazdin, 1981).

Researchers have generally targeted potential mediators of interventions, especially staff in developmental disabilities services, to gather data on social validity. Existing studies have focused on a range of factors that may affect staff views on social validity, but the most reliable effects have proved to be the type of intervention, and severity of the presenting problem (Elliott, 1988; Lennox & Miltenberger, 1990; Miltenberger, 1990; Reimers, Wacker, & Koeppl, 1987; Storey & Horner, 1991). Specifically, reinforcement-based interventions (e.g., differential reinforcement) and interventions that are relatively less intrusive are rated as more acceptable, and the acceptability of intrusive interventions increases as problem severity increases.

A more important question for the present discussion is whether interventions based on functional assessment are perceived as more socially acceptable by consumers. Research studies published to date cannot be used to draw a strong conclusion with one study finding that interventions linked to functional assessment are rated as more socially valid (Jones & Lungaro, 2000), two studies finding weak or inconsistent effects (Hastings, Boulton, Monzani, & Tombs, 2004; Weigle & Scotti, 2000), and one study finding no effects (Miltenberger & Lumley, 1997).

Further Issues

In addition to the main questions reviewed above, there are some further ethical issues that have a bearing on the balance of the debate. Those discussed here relate to: 1. The time and personnel investment in conducting functional assessment, especially analog assessment, 2. The use of preventative interventions, 3. The elicitation of self-injury during analog assessment procedures, and 4. The clarity of results from functional assessment methods. Each of these is discussed in turn below.

Functional assessment takes time and specialist resources to carry out. In the case of a full analog assessment, data collection takes hours over several days and also requires considerable training in observational skills and data presentation and interpretation (see Iwata et al., 1982). Even for non-experimental methods, an experienced behavior analyst must spend considerable time carrying out analyses and the time input from other service staff to provide data for a functional assessment is significant. Although brief methods of analysis have been developed (e.g., Kahng & Iwata, 1999; Northup et al., 1991), these still require an expert team and are often conducted in a specialized outpatient setting. The general question is whether time and resources would be best employed simply teaching clients key skills such as functional communication techniques. Given that around 70% of functional assessments seem to identify either attention or demand avoidance functions for challenging behaviors including self-injury (Derby et al., 1992; Iwata et al., 1994), appropriate behaviors for achieving these two functions could perhaps be taught...
during the time that it takes to conduct functional assessments.

This last point also relates to a broader question about preventative intervention. Self-injurious behaviors do not appear from nowhere, but are shaped over time by exposure to environmental and especially social contingencies (Guess & Carr, 1991). This developmental process can be understood partly through the concept of the establishing operation (Michael, 1982). Michael promoted this concept as a way of offering a behavior analytic explanation of motivation. Fundamentally, the motivation for an individual to obtain positive reinforcement occurs as a direct consequence of a period of deprivation of that reinforcement. Similarly, the motivation for negative reinforcement occurs as a direct consequence of exposure to an aversive event. Applying this notion to the development of self-injury, these behaviors may function to overcome a state of deprivation (such as a period of no social contact or activity), or to avoid an aversive situation (such as an academic demand). Without the environmental conditions to establish self-injury as reinforcing, the behavior would not occur.

Thus, getting to the stage of the need for a functional assessment is in many ways too late. Intervening early to identify establishing operations that increase likelihood of self-injury occurring is an obvious therapeutic goal. It may involve preventing states of deprivation (e.g., lack of social attention), through techniques such as non-contingent reinforcement (Vollmer, Iwata, Zarcone, Smith, & Mazaleski, 1993), to enabling more adaptive means of avoiding demands or gaining attention through teaching functional communication skills (Dunlap, Foster-Johnson, & Robbins, 1990; Reeve & Carr, 2000). Ultimately, such preventative strategies may prove to be more efficient and effective (and thus ethical) treatment approaches.

The preceding discussions about intervening either without a functional assessment or earlier in the developmental history of self-injurious behaviors are important points relating to the merits of conducting functional assessments versus not doing so. However, there is a broader issue about these interventions. We characterized behavior modification as the application of behavior change procedures without a functional assessment, and a strict interpretation of this definition might encompass interventions designed to teach skills such as functional communication without an explicit functional assessment. However, the interventions in these cases would arguably be based on an implicit functional assessment built on the basis of evidence across a large number of published cases. There is also another sense in which these general preventative interventions cannot be considered as simply behavior modification and that is the fact that they are clearly constructional (cf. Goldiamond, 1974) in nature.

A further issue is that experimental analog assessment is designed to elicit self-injury under controlled experimental conditions. Given that one would want to avoid unnecessary harm to clients, it is important to consider that self-injury can be seen to be ‘made’ to occur within analog assessment purely to enable hypotheses about its function to be derived. Thus, behavior analysts would need to be convinced of the validity of analog assessment methods in balancing the ethical case for their use. However, there is a range of internal and external validity problems with the experimental conditions that are a constituent part of analog assessment (Oliver, 1991; Sturmey, 1995). For example, there is a possibility that self-injury could be shaped within the analyses by contingencies that have not in fact been responsible for maintaining the behavior in the natural environment. Therefore, the ethical case for use of analog assessment is not watertight. This particular criticism does not necessarily extend to all functional assessment methods.

A final ethical point to be discussed here is that outcomes of functional assessments are not always clear. In particular, no clear function or multiple functions might be identified in a significant minority of cases (Derby et al., 1992; Iwata et al., 1994). When such results are obtained, the links to intervention are much less clear and it is then questionable whether the effort in conducting a functional assessment has been worthwhile. Of course, there is a balance of risk here in that for many cases of self-injury it is possible to identify clear functional hypotheses and that these may be the optimal conditions for a maximally
effective and thus ethically superior treatment.

Conclusions

The present review suggests that basing treatment of self-injury on results of a prior functional assessment is more ethical than an eliminative behavior modification approach. Firstly, treatment based on functional analysis is more effective overall, more likely to generalize and maintain, and is probably more likely to be implemented with good fidelity by staff (presumably increasing the chance of successful treatment). Secondly, to the extent that treatments based on functional assessments are more likely to be constructional in nature they are also likely to be less restrictive than behavior modification approaches. Thirdly, although direct data from the small number of empirical studies are ambiguous, to the extent that procedures based on functional assessment use reinforcement-based strategies they are likely to be viewed by consumers as more acceptable.

These conclusions lend support to use of functional assessments in the treatment of self-injury. However, it is also important to consider limitations of functional assessment approaches. These include the intensity of resources required, a potential lack of clarity of results, and questions about validity. Furthermore, there may be specific situations where self-injury is immediately life-threatening when a behavior modification procedure is a swift and ethical intervention. These factors and others may limit the ethical superiority of treatments based on functional assessment and indicate issues to be addressed in the further development of functional assessment and functional interventions for self-injurious behavior.

References


Hastings, R. P. (2002). Do challenging behaviors affect staff psychological well-being?: Issues of...


lytic approach (pp. 139–163). Chichester, UK: Wiley.

Received: 21 April 2004
Initial Acceptance: 30 June 2004
Final Acceptance: 1 December 2004